

# Groups with poly-context-free word problem

Tara Brough

April 10, 2011

## Abstract

We consider the class of groups whose word problem is *poly-context-free*; that is, an intersection of finitely many context-free languages.

We show that any group which is virtually a finitely generated subgroup of a direct product of free groups has poly-context-free word problem, and conjecture that the converse also holds. We prove our conjecture for several classes of soluble groups, including metabelian groups and torsion-free soluble groups, and present progress towards resolving the conjecture for soluble groups in general.

Some of the techniques introduced for proving languages not to be poly-context-free may be of independent interest.

## 1 Introduction

The *word problem* of a group  $G$  with respect to a finite generating set  $X$ , denoted  $W(G, X)$ , is the set of all words in elements of  $X$  and their inverses which represent the identity element of  $G$ . A *(formal) language* is a set of words over some finite alphabet, so  $W(G, X)$  can be considered as a language.

The study of word problems of groups as languages has developed slowly since the beginnings of language theory in the 1950s. In 1971, Anisimov [1] published a proof that a group has regular word problem if and only if it is finite. The first really significant development in the area was the classification of the groups with context-free word problem by Muller and Schupp in the 1980s [23, 24, 4]: a finitely generated group has context-free word problem if and only if it is virtually free. Since then, research activity

in this area has increased, and groups with word problem in various other language classes, generally somewhat related to the context-free languages, have been studied, for example in [10, 11, 12, 13, 14, 15, 16, 19, 20, 26]. The general aim is to determine what implications the language type of a group's word problem has for the structure of the group and vice versa.

One natural class of languages to consider is the closure of the context-free languages under intersection. Some research has been done on this class (see for example [21], [27] and [9]), but it does not appear to have a consistent name. We call a language *k-context-free* (henceforth abbreviated to *k-CF*) if it is an intersection of finitely many context-free languages, and *poly-context-free* (*poly-CF*) if it is *k-CF* for some  $k \in \mathbb{N}$ .

This paper is concerned with the class of *poly-CF* groups. A group is said to be *poly-CF* if its word problem is a *poly-CF* language. The property of being *poly-CF* is independent of the choice of finite generating set, and the class of *poly-CF* groups is closed under taking finitely generated subgroups, finite index overgroups, and finite direct products. All but the last of these properties are typical of classes of groups defined by the language type of their word problem.

A general classification of these groups appears to be hard. However, we prove a result (Theorem 5.2) which comes close in the case of soluble groups.

We conjecture that the only *poly-CF* groups are those obtained from virtually free groups using the above-mentioned operations; that is, that a group is *poly-CF* if and only if it is virtually a finitely generated subgroup of a direct product of free groups (Conjecture 5.1). This would mean that the only soluble *poly-CF* groups are the virtually abelian groups. Theorem 5.2 gives substantial evidence towards this special case of our conjecture.

In [11], the *coCF* groups (groups whose word problem is the complement of a context-free language) were studied. Various closure properties of the *coCF* groups were determined, most of which carry over easily to *poly-CF* groups (see Proposition 2.5 below). Additionally, several classes of groups were shown not to be *coCF*, using a method [11, Proposition 14] based on the correspondence between context-free languages and semilinear sets (see Section 2.4.1 below). We prove a strengthened version of [11, Proposition 14] (see Proposition 3.2 below), which enables us to deduce that any group proved not *coCF* using [11, Proposition 14] is also not *poly-CF*. Some examples of such groups are finitely generated nilpotent or polycyclic groups

that are not virtually abelian. It was these results that led to the attempt at a characterisation of the soluble poly- $\mathcal{CF}$  groups.

A major open problem for  $\text{co}\mathcal{CF}$  groups is whether they are closed under taking free products. It was suggested by Derek Holt that closure under free products might be much easier to determine for poly- $\mathcal{CF}$  groups, but so far this problem also remains open, though we believe that the word problem of  $\mathbb{Z}^2 * \mathbb{Z}$  is not poly- $\mathcal{CF}$ . The poly- $\mathcal{CF}$  groups are somewhat related to the  $\text{co}\mathcal{CF}$  groups, in the sense that if our main conjecture is true, then the poly- $\mathcal{CF}$  groups are a subclass of the  $\text{co}\mathcal{CF}$ -groups, as we explain in Section 4, following Conjecture 4.2.

Our main tools are introduced in Section 2. These are: various closure properties of the classes of poly- $\mathcal{CF}$  languages and poly- $\mathcal{CF}$  groups; the relationship between bounded context-free languages and semilinear sets, due to Parikh [25] and Ginsburg and Spanier [8]; and a result by the author and Derek Holt [3], showing that every finitely generated soluble group that is not virtually abelian has a subgroup isomorphic to one of a small number of types.

In Section 3, we study the class of poly- $\mathcal{CF}$  languages, with a particular focus on methods for proving languages to be *not* poly- $\mathcal{CF}$ . To this end, we develop several tools based on the correspondence between context-free languages and stratified semilinear sets introduced in Section 2.4.1. In Corollary 3.3, we show that a language satisfying certain properties is neither poly- $\mathcal{CF}$  nor  $\text{co}\mathcal{CF}$ , while Theorem 3.12 exhibits sequences of languages  $L^{(n,k)}$ , where  $n, k \in \mathbb{N}$ , such that for all  $n$ , the language  $L^{(n,k)}$  is an intersection of  $k$  but not  $k - 1$  context-free languages. This is an extension of a result by Liu and Weiner [21].

In Section 4, we present the known examples of poly- $\mathcal{CF}$  groups, and conjecture that these are the only ones. We give some evidence for this conjecture (Conjecture 4.2), in the form of results showing that it holds in the classes of nilpotent, Baumslag-Solitar and polycyclic groups, and for the groups  $G(\mathbf{c})$  introduced in [3], which are also shown to be not  $\text{co}\mathcal{CF}$  if they are not virtually abelian.

We conclude with a section applying the results of Section 4 and [3] to prove the metabelian and torsion-free soluble cases of our conjecture, and to narrow down the possibilities for which soluble groups could be poly- $\mathcal{CF}$ .

## 2 Background and notation

### 2.1 Notation

$\mathbb{N}, \mathbb{Z}$  and  $\mathbb{Q}$  denote the natural numbers, integers and rationals respectively. We denote the natural numbers with zero included by  $\mathbb{N}_0$ .

For  $r \in \mathbb{N}$  and  $1 \leq i \leq r$ , the vector in  $\mathbb{N}_0^r$  with a 1 in the  $i$ -th position and zeroes elsewhere will be denoted by  $e_i$ . With the exception of these, all vectors are represented by bold letters. We denote the  $i$ -th component of the vector  $\mathbf{v}$  by  $\mathbf{v}(i)$ .

For a set  $X$ , we denote the *Kleene star closure* of  $X$ , which is the set of all finite length strings (also called words) of elements of  $X$ , by  $X^*$ . In the special case  $X = \{x\}$ , we often denote  $X^*$  by  $x^*$ .

### 2.2 Closure properties of the poly- $\mathcal{CF}$ languages

Many closure properties of the classes of  $k\text{-}\mathcal{CF}$  and poly- $\mathcal{CF}$  languages can be deduced from the similar properties for context-free languages; for details of these, see (for example) [17].

**Proposition 2.1.** *For any  $k \in \mathbb{N}$ , the class of  $k\text{-}\mathcal{CF}$  languages is closed under inverse homomorphisms, inverse generalised sequential machine mappings, union with context-free languages and intersection with regular languages. The class of poly- $\mathcal{CF}$  languages is closed under all these operations, and also under intersection and union.*

*Proof.* Let  $L = L_1 \cap \dots \cap L_k$  with each  $L_i$  context-free and let  $\Sigma$  be the alphabet of  $L$ . Let  $\Gamma$  be an alphabet and let  $\phi$  be a homomorphism from  $\Gamma^*$  to  $\Sigma^*$ , or a generalised sequential machine mapping with input alphabet  $\Gamma$  and output alphabet  $\Sigma$ . Then

$$\begin{aligned} \phi^{-1}(L) &= \{w \in \Gamma^* \mid \phi(w) \in L_i \ (1 \leq i \leq k)\} \\ &= \bigcap_{i=1}^k \{w \in \Gamma^* \mid \phi(w) \in L_i\} = \bigcap_{i=1}^k \phi^{-1}(L_i), \end{aligned}$$

and so, since the class of context-free languages is closed under inverse homomorphisms and inverse generalised sequential machine mappings,  $\phi^{-1}(L)$

is  $k\text{-}\mathcal{CF}$ .

The class of context-free languages is closed under union and under intersection with regular languages. Thus if  $R$  is regular, then  $L \cap R = L_1 \cap \dots \cap L_{k-1} \cap (L_k \cap R)$  is  $k\text{-}\mathcal{CF}$ ; and if  $M$  is context-free, then  $L \cup M = \bigcap_{i=1}^k (L_i \cup M)$  is  $k\text{-}\mathcal{CF}$ .

The closure of the class of poly- $\mathcal{CF}$  languages under intersection is obvious, since if  $L_1$  is  $k_1\text{-}\mathcal{CF}$  and  $L_2$  is  $k_2\text{-}\mathcal{CF}$ , then  $L_1 \cap L_2$  is an intersection of  $k_1 + k_2$  context-free languages.

If  $L = \bigcap_{i=1}^m L_i$  and  $M = \bigcap_{j=1}^n M_j$ , with each  $L_i$  and  $M_j$  context-free, then

$$L \cup M = \left( \bigcap_{i=1}^m L_i \right) \cup \left( \bigcap_{j=1}^n M_j \right) = \bigcap_{i=1}^m \bigcap_{j=1}^n (L_i \cup M_j)$$

is  $mn\text{-}\mathcal{CF}$ , so the class of poly- $\mathcal{CF}$  languages is also closed under union.  $\square$

The closure of the poly- $\mathcal{CF}$  languages under union and intersection was already observed by Wotschke [27], who also showed, using a theorem of Liu and Weiner (see Section 3.2 below), that the poly- $\mathcal{CF}$  languages are not closed under complementation and are thus properly contained in the Boolean closure of the context-free languages [27, Theorem II.4].

Any recursively enumerable language can be expressed as a homomorphic image of the intersection of two deterministic context-free languages [5]. Every poly- $\mathcal{CF}$  language is context-sensitive, since the context-sensitive languages are closed under intersection and contain the context-free languages. Thus the poly- $\mathcal{CF}$  languages are not closed under homomorphisms.

### 2.3 Basic properties of the poly- $\mathcal{CF}$ groups

A central result in the theory of word problems of groups as languages is the following, for which a proof is given in [11]. We denote the complement of  $W(G, X)$  in  $X^*$  by  $\text{co}W(G, X)$ .

**Lemma 2.2.** [11, Lemma 1] *Let  $\mathcal{C}$  be a class of languages closed under inverse homomorphisms and let  $G$  be a finitely generated group. Then the following hold.*

- (i)  $W(G, X) \in \mathcal{C}$  for some finite generating set  $X$  if and only if for every finite generating set  $Y$ ,  $W(G, Y) \in \mathcal{C}$ .
- (ii)  $\text{co}W(G, X) \in \mathcal{C}$  for some finite generating set  $X$  if and only if for every finite generating set  $Y$ ,  $\text{co}W(G, Y) \in \mathcal{C}$ .

In this case, we call  $G$  a  $\mathcal{C}$  group if  $W(G)$  is in  $\mathcal{C}$ , and a  $\text{co}\mathcal{C}$  group if  $\text{co}W(G)$  is in  $\mathcal{C}$ , and say that  $\mathcal{C}$  groups or  $\text{co}\mathcal{C}$  groups are *insensitive* to choice of generators.

**Lemma 2.3.** [11, Lemma 2] *Let  $\mathcal{C}$  be a class of languages closed under inverse homomorphisms and intersection with regular sets. Then the classes of  $\mathcal{C}$  groups and  $\text{co}\mathcal{C}$  groups are closed under taking finitely generated subgroups.*

**Lemma 2.4.** [11, Lemma 5] *Let  $\mathcal{C}$  be a class of languages closed under union with regular sets and inverse generalised sequential machine mappings. Then the classes of  $\mathcal{C}$  groups and  $\text{co}\mathcal{C}$  groups are closed under passing to finite index overgroups.*

Thus, by Proposition 2.1, we have:

**Proposition 2.5.** *The classes  $\text{co}\mathcal{CF}$  and  $k\text{-}\mathcal{CF}$  groups (for any  $k \in \mathbb{N}$ ) are insensitive to choice of generators and closed under passing to finitely generated subgroups and passing to finite index overgroups.*

## 2.4 Semilinear sets

A useful tool for proving languages not to be poly- $\mathcal{CF}$  is a relationship between context-free languages and semilinear sets, introduced by Parikh [25] and then strengthened, in the case of bounded languages, by Ginsburg and Spanier [8].

A *linear set* is a subset  $L$  of  $\mathbb{N}_0^r$  for which there exist a *constant vector*  $\mathbf{c} \in \mathbb{N}_0^r$  and a finite set of *periods*  $P = \{\mathbf{p}_i \mid 1 \leq i \leq n\} \subseteq \mathbb{N}_0^r$  such that

$$L = \{\mathbf{c} + \sum_{i=1}^n \alpha_i \mathbf{p}_i \mid \alpha_i \in \mathbb{N}_0\}.$$

Note that the set of periods  $P$  is not uniquely determined. A *semilinear set* is a union of finitely many linear sets.

Following Ginsburg [6], we will use the notation  $L(\mathbf{c}; \mathbf{p}_1, \dots, \mathbf{p}_n)$ , or  $L(\mathbf{c}; P)$ , for a linear set with constant  $\mathbf{c}$  and set of periods  $P = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ . For  $C$  a set of constant vectors, we will denote  $\bigcup_{\mathbf{c} \in C} L(\mathbf{c}; P)$  by  $L(C; P)$ . If  $C = \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$ , we will also write  $L(\mathbf{c}_1, \dots, \mathbf{c}_m; \mathbf{p}_1, \dots, \mathbf{p}_n)$  for  $L(C; P)$ .

If  $L = L(\mathbf{c}; P)$ , we define  $L^{\mathbb{Q}}$  to be the set  $\{\mathbf{c} + \sum_{i=1}^n a_i \mathbf{p}_i \mid a_i \in \mathbb{Q}\}$ . This is a coset in  $\mathbb{Q}^n$  of the  $\mathbb{Q}$ -subspace spanned by  $P$ . We define  $L^{\mathbf{0}}$  to be  $L(\mathbf{0}; P)$ , that is, the linear set having the same periods as  $L$  and constant  $\mathbf{0}$ .

A subset  $P$  of  $\mathbb{N}_0^r$  is *stratified* if it satisfies the following conditions:

- (i) each  $\mathbf{p} \in P$  has at most two non-zero components, and
- (ii) there do not exist  $i < j < k < l$  and non-zero  $a, b, c, d \in \mathbb{N}$  such that  $ae_i + be_k$  and  $ce_j + de_l$  are both in  $P$ .

A linear set is *stratified* if it can be expressed using a stratified set of periods. A semilinear set is *stratified* if it can be expressed as a union of finitely many stratified linear sets. (We follow Liu and Weiner [21] for this terminology.) Note that stratified linear and semilinear sets are not generally stratified sets in the sense of the previous paragraph.

#### 2.4.1 Stratified semilinear sets and bounded poly- $\mathcal{CF}$ languages

The *commutative image* of a language  $L$  over  $\{a_1, \dots, a_r\}$  is the subset of  $\mathbb{N}_0^r$  given by mapping each  $w \in L$  to the tuple  $(n_1, \dots, n_r)$ , where  $n_i$  is the number of occurrences of  $a_i$  in  $w$ . Parikh's theorem [25] says that the commutative image of a context-free language is always a semilinear set. The converse of Parikh's theorem does not hold: consider for example the language  $\{a^m b^n c^m d^n \mid m, n \in \mathbb{N}_0\}$ .

A language  $L \subseteq X^*$  is *bounded* if there exist  $w_1, \dots, w_n \in X^*$  such that  $L \subseteq w_1^* \dots w_n^*$ , in which case we can define a corresponding subset of  $\mathbb{N}_0^n$ :

$$\Phi(L) = \{(m_1, \dots, m_n) \mid m_i \in \mathbb{N}_0, w_1^{m_1} \dots w_n^{m_n} \in L\}.$$

When  $w_1, \dots, w_n$  are distinct single symbols, this is the same as the commutative image of  $L$ . Thus the following result of Ginsburg and Spanier strengthens Parikh's theorem in the case of bounded languages.

**Theorem 2.6.** [6, Theorem 5.4.2] *Let  $W \subseteq w_1^* \dots w_n^*$ , each  $w_i$  a word. Then  $W$  is context-free if and only if  $\Phi(W)$  is a stratified semilinear set.*

Ginsburg and Spanier used different notation, which made it more transparent how to get from  $\Phi(W)$  back to  $W$ . But as we will only require the ‘only if’ direction, we prefer this tidier notation.

Theorem 2.6 is easily extended to the poly- $\mathcal{CF}$  languages.

**Corollary 2.7.** *If  $L$  is a  $k$ - $\mathcal{CF}$  language, then for any  $w_1, \dots, w_n$ , the subset  $\Phi(L \cap w_1^* \dots w_n^*)$  of  $\mathbb{N}_0^n$  is an intersection of  $k$  stratified semilinear sets.*

*Proof.* Let  $L = L_1 \cap \dots \cap L_k$  with each  $L_i$  context-free, and let  $W = w_1^* \dots w_n^*$ , where each  $w_i$  is a word in the alphabet of  $L$ . For  $1 \leq i \leq k$ , let  $M_i = L_i \cap W$ . Then  $L \cap W = L_1 \cap \dots \cap L_k \cap W = \bigcap_{i=1}^k M_i$  and

$$\begin{aligned} \Phi(L \cap W) &= \{(m_1, \dots, m_n) \mid m_i \in \mathbb{N}_0, w_1^{m_1} \dots w_n^{m_n} \in L \cap W\} \\ &= \bigcap_{i=1}^k \{(m_1, \dots, m_n) \mid m_i \in \mathbb{N}_0, w_1^{m_1} \dots w_n^{m_n} \in M_i\} \\ &= \bigcap_{i=1}^k \Phi(M_i) \end{aligned}$$

and each  $\Phi(M_i)$  is a stratified semilinear set by Theorem 2.6.  $\square$

Proving that a given semilinear set is not stratified is by no means straightforward, since there can be many different ways of expressing a semilinear set as a union of finitely many linear sets. Ginsburg [6] mentioned that there was no known decision procedure for determining whether an arbitrary semilinear set is stratified, and it appears that this is still an open problem.

#### 2.4.2 Closure properties of the class of semilinear sets

The class of semilinear sets is obviously closed under union. Thinking geometrically, one would also expect this class to be closed under the other Boolean operations. This is indeed true, but much less easy to show.



The intersection of finitely many linear sets is always a semilinear set of quite a restricted form. This result can be derived from the proof of Theorem 5.6.1 in [6]. For another proof, obtained independently by the author, see [2, Proposition 2.3].

**Proposition 2.8.** *If  $L$  is the nonempty intersection of linear subsets  $L_1, \dots, L_n$  of  $\mathbb{N}_0^r$ , then  $L$  is semilinear. Moreover,*

$$L = L(\mathbf{C}_1, \dots, \mathbf{C}_k; \mathbf{P}_1, \dots, \mathbf{P}_m),$$

*where  $\mathbf{C}_i \in \mathbb{N}_0^r$ , and  $\mathbf{P}_1, \dots, \mathbf{P}_m$  are such that  $\bigcap_{i=1}^n L_i^0 = L(\mathbf{0}; \mathbf{P}_1, \dots, \mathbf{P}_m)$ . If  $L_1, \dots, L_n$  all have constant vector zero, then  $L$  is linear with constant vector zero.*

**Corollary 2.9.** [6, Theorem 5.6.1] *Let  $L$  be an intersection of finitely many semilinear sets. Then  $L$  is a semilinear set.*

**Proposition 2.10.** [7, Theorem 6.2 and Corollary 1] *If  $L$  and  $M$  are semilinear subsets of  $\mathbb{N}_0^r$ , then  $M - L$  is also a semilinear subset of  $\mathbb{N}_0^r$  and effectively calculable from  $L$  and  $M$ . In particular, since  $\mathbb{N}_0^r$  is semilinear, if  $L$  is a semilinear subset of  $\mathbb{N}_0^r$ , then the complement of  $L$  in  $\mathbb{N}_0^r$  is semilinear.*

### 2.4.3 Dimension of linear sets

If  $V$  is a subspace of a vector space  $W$  with  $\dim(V) < \dim(W)$ , then the *dimension* of a coset of  $V$  in  $W$  is defined to be the dimension of  $V$ . The *dimension* of a linear set  $L$  is defined to be the dimension of  $L^{\mathbb{Q}}$  or, equivalently, the dimension of the vector space over  $\mathbb{Q}$  spanned by the periods of  $L$ .

We record here a result about the dimension of linear sets which will be useful later. This is a known result, but the only reference we have for it is [21], where the proof given is incorrect. A proof is included in the author's Ph.D. thesis [2, Proposition 2.10].

**Proposition 2.11.** *A linear set of dimension  $n + 1$  cannot be expressed as a union of finitely many linear sets of dimension  $n$  or less.*

## 2.5 Subgroups of finitely generated soluble groups

The following theorem is a combination of Theorems 3.3 and 5.2 in [3]. By  $\mathbb{Z}^\infty$ , we mean the free abelian group of countably infinite rank. For the

definition of a proper Gc-group, see Section 4.3. A group is *metabelian* if it has derived length at most 2.

**Theorem 2.12.** *Let  $G$  be a finitely generated soluble group which is not virtually abelian. Then  $G$  has a subgroup isomorphic to at least one of the following.*

- (i)  $\mathbb{Z}^\infty$ ;
- (ii) a proper Gc-group;
- (iii) a finitely generated group  $H$  with an infinite normal torsion subgroup  $U$ , such that  $H/U$  is either free abelian or a proper Gc-group.

*If  $G$  is metabelian, then the subgroup  $H$  in (iii) can always be taken to be  $C_p \wr \mathbb{Z}$  for some prime  $p$ .*

Since the class of poly- $\mathcal{CF}$  groups is closed under taking finitely generated subgroups, this gives a very useful approach towards resolving our conjecture for soluble groups.

### 3 Poly- $\mathcal{CF}$ languages

Recall that a  $k$ - $\mathcal{CF}$  language is an intersection of  $k$  context-free languages, and a poly- $\mathcal{CF}$  language is a language which is  $k$ - $\mathcal{CF}$  for some  $k \in \mathbb{N}$ . In this section, we shall primarily be concerned with proving some results which will assist us in determining that the word problems of certain groups are not poly- $\mathcal{CF}$ .

#### 3.1 A criterion for a language to be neither poly- $\mathcal{CF}$ nor $\text{co}\mathcal{CF}$

In [11, Proposition 14], a technique was developed for proving a subset of  $\mathbb{N}_0^r$  not to be the complement of a semilinear set. This was used in combination with Parikh's theorem to prove various classes of groups not to be  $\text{co}\mathcal{CF}$ . The proof is fairly long and technical. The authors were presumably unaware of the fact that the complement of a semilinear set is semilinear. This fact allows us to give a much simpler proof of their result, and to strengthen it.

If  $\mathbf{a}$  and  $\mathbf{b}$  are vectors in  $\mathbb{N}_0^r$  and  $\mathbb{N}_0^s$  respectively, then we denote by  $(\mathbf{a}; \mathbf{b})$  the vector in  $\mathbb{N}_0^{r+s}$  which consists of all the components of  $\mathbf{a}$  in order, followed by those of  $\mathbf{b}$  in order. When talking about vectors in  $\mathbb{N}_0^{r+s}$ , if we write  $(\mathbf{a}; \mathbf{b})$ , then it is understood that  $\mathbf{a} \in \mathbb{N}_0^r$  and  $\mathbf{b} \in \mathbb{N}_0^s$ . For  $\mathbf{a} \in \mathbb{N}_0^r$ , we define  $\sigma(\mathbf{a}) = \sum_{i=1}^r \mathbf{a}(i)$ .

We use the following lemma, extracted from the proof of Proposition 11 in [11]. We call a vector  $\mathbf{v} \in \mathbb{N}_0^{r+s}$  *simple* if its first  $r$  components are all zero, and *complex* otherwise. The proof is quoted from [11] with only minor modifications.

**Lemma 3.1.** *Let  $L = L_1 \cup \dots \cup L_n$ , with each  $L_i$  a linear subset of  $\mathbb{N}_0^{r+s}$ . Then there exists a constant  $C \in \mathbb{N}$  such that if  $(\mathbf{a}; \mathbf{b}) \in L$  can be expressed using only complex periods, then  $\mathbf{b}(j) < C\sigma(\mathbf{a})$  for all  $1 \leq j \leq s$ .*

*Proof.* Fix some  $i \in \{1, \dots, n\}$  and let  $L_i = L(\mathbf{c}_i; P_i)$ . If  $(\mathbf{p}; \mathbf{q}) \in P_i$  is a complex period, then  $\sigma(\mathbf{p}) \neq 0$ , so there exists  $t$  such that  $\mathbf{q}(j) < t\sigma(\mathbf{p})$  for  $1 \leq j \leq s$ . Since  $P_i$  is finite, we can choose the same  $t$  for all  $(\mathbf{p}; \mathbf{q}) \in P_i$ . If, for  $k = 1, 2$ ,  $(\mathbf{a}_k; \mathbf{b}_k) \in \mathbb{N}_0^{r+s}$  satisfy  $\mathbf{b}_k(j) < t\sigma(\mathbf{a}_k)$ , then  $(\mathbf{b}_1 + \mathbf{b}_2)(j) < t\sigma(\mathbf{a}_1 + \mathbf{a}_2)$ . Thus there is a constant  $q \in \mathbb{N}_0$ , which can be taken to be  $\max\{\mathbf{c}_i(j) \mid 1 \leq j \leq r\}$ , such that if  $(\mathbf{a}; \mathbf{b}) \in L_i$  can be expressed using only complex periods, then  $\mathbf{b}(j) < t\sigma(\mathbf{a}) + q$  for all  $1 \leq j \leq s$ .

Now let  $C \in \mathbb{N}$  be twice the maximum of all of the constants  $t, q$  that arise for all  $L_i$ . Then, for any  $(\mathbf{a}; \mathbf{b}) \in L$  which can be expressed using only complex periods,  $\mathbf{b}(j) < C\sigma(\mathbf{a})$  for all  $1 \leq j \leq s$ .  $\square$

**Proposition 3.2.** *Let  $L \subseteq \mathbb{N}_0^{r+s}$  for some  $r, s \in \mathbb{N}$ . Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be an unbounded function and suppose that, for every  $k \in \mathbb{N}$ , there exists  $\mathbf{a} \in \mathbb{N}_0^r \setminus \{\mathbf{0}\}$  such that the following hold:*

- (i) *There exists  $\mathbf{b} \in \mathbb{N}_0^s$  such that  $(\mathbf{a}; \mathbf{b}) \in L$ .*
- (ii) *If  $(\mathbf{a}; \mathbf{b}) \in L$  then  $\mathbf{b}(j) \geq k\sigma(\mathbf{a})$  for some  $1 \leq j \leq s$ .*
- (iii) *If  $(\mathbf{a}; \mathbf{b}), (\mathbf{a}; \mathbf{b}') \in L$  with  $\mathbf{b} \neq \mathbf{b}'$ , then  $|\mathbf{b}(l) - \mathbf{b}'(l)| \geq f(k)$  for some  $1 \leq l \leq s$ .*

*Then  $L$  is not a semilinear set.*

*Proof.* Let  $L$  be as in the statement of the proposition and suppose that  $L = \bigcup_{i=1}^n L_i$ , where each  $L_i = L(\mathbf{c}_i; P_i)$  is a linear subset of  $\mathbb{N}_0^{r+s}$ . By Lemma 3.1, there exists a constant  $C \in \mathbb{N}$  such that if  $(\mathbf{a}; \mathbf{b}) \in L$  can be expressed using only complex periods in some  $L_i$ , then  $\mathbf{b}(j) < C\sigma(\mathbf{a})$  for all  $1 \leq j \leq s$ .

Choose  $k > C$ , and suppose  $\mathbf{a}$  satisfies the hypotheses of the proposition with respect to  $k$ . If  $(\mathbf{a}; \mathbf{b}) \in L$ , then  $(\mathbf{a}; \mathbf{b})$  cannot be expressed using only complex periods, so some  $P_i$  must contain a simple period  $(\mathbf{0}; \mathbf{v})$  with  $\mathbf{v}$  non-zero. But then  $(\mathbf{a}; \mathbf{b} + \mathbf{v}) \in L_i \subseteq L$  and so, for some  $1 \leq l \leq s$ ,

$$|\mathbf{v}(l)| = |(\mathbf{b} + \mathbf{v})(l) - \mathbf{b}(l)| \geq f(k).$$

So for all  $k > C$ , there is a non-zero simple period  $\mathbf{v}_k$  in  $\bigcup_{i=1}^n P_i$ , with some component of  $\mathbf{v}_k$  being at least  $f(k)$ . But since  $\bigcup_{i=1}^n P_i$  is finite and  $f(k)$  is unbounded, this is impossible. Thus  $L$  is not a semilinear set.  $\square$

In [11, Proposition 14], instead of our condition (i), it is required that there is a *unique*  $\mathbf{b} \in \mathbb{N}_0^r$  such that  $(\mathbf{a}; \mathbf{b}) \in L$  (and thus there is no condition (iii) or mention of the unbounded function  $f$ ); instead of our condition (ii), it is required that  $\mathbf{b}(j) \geq k\sigma(\mathbf{a})$  for *every*  $1 \leq j \leq s$ . The conclusion is that  $L$  is not the complement of a semilinear set.

Our hypothesis is considerably weaker, and the conclusion is equally strong, since the complement of a semilinear set is semilinear.

For  $\mathbf{v} = (n_1, \dots, n_r) \in \mathbb{N}_0^r$  and  $\tau$  a permutation of  $\{1, \dots, r\}$ , we define

$$\tau(\mathbf{v}) = (n_{\tau(1)}, n_{\tau(2)}, \dots, n_{\tau(r)}).$$

We extend this to a subset  $L$  of  $\mathbb{N}_0^r$  by defining  $\tau(L) = \{\tau(\mathbf{v}) \mid \mathbf{v} \in L\}$ . If  $L = L(\mathbf{c}; \mathbf{p}_1, \dots, \mathbf{p}_k)$ , then  $\tau(L) = (\tau(\mathbf{c}); \tau(\mathbf{p}_1), \dots, \tau(\mathbf{p}_k))$ , so the property of being a linear set, or indeed an intersection of  $k$  semilinear sets, is preserved by  $\tau$ .

We shall make significant use of the following corollary to Proposition 3.2 in Section 4.

**Corollary 3.3.** *Let  $L \subseteq w_1^* \dots w_k^*$  be a bounded language over an alphabet  $X$  with  $w_i \in X^*$ , and let  $\tau$  be a permutation of  $\{1, \dots, k\}$ . If  $\tau(\Phi(L))$  satisfies the hypothesis of Proposition 3.2, then  $L$  is neither  $\text{coCF}$  nor  $\text{poly-CF}$ .*

*Proof.* Since  $\tau$  preserves semilinearity, this follows immediately from Proposition 3.2, Theorem 2.6 and Corollary 2.7 and the fact that the class of semilinear sets is closed under intersection (Corollary 2.9) and complementation (Proposition 2.10).  $\square$

### 3.2 The languages $L^{(k)}$

A  $(k - 1)$ - $\mathcal{CF}$  language is clearly also  $n$ - $\mathcal{CF}$  for all  $n \geq k$ . In [21], Liu and Weiner showed that the class of  $k$ - $\mathcal{CF}$  languages properly contains the class of  $(k - 1)$ - $\mathcal{CF}$  languages, thus exhibiting an infinite hierarchy of languages in between the context-free and context-sensitive languages. (They call a  $k$ - $\mathcal{CF}$  language a ‘ $k$ -intersection language’.) Note that this implies that the  $k$ - $\mathcal{CF}$  languages are not closed under intersection or even under intersection with context-free languages.

There are some problems with Liu and Weiner’s proof, particularly in the proof of their Theorem 10. In this section, we provide a more detailed proof. In Section 3.3, we extend Liu and Weiner’s result, but the proof of the special case is provided first, as it will probably aid the reader’s understanding of the more general case.

Following Liu and Weiner, we define a sequence of languages  $L^{(k)}$  and corresponding subsets  $S^{(k)}$  of  $\mathbb{N}_0^{2k}$ . For  $k \in \mathbb{N}$ , let  $a_1, \dots, a_{2k}$  be  $2k$  distinct symbols, and define the language

$$L^{(k)} = \{a_1^{n_1} \dots a_k^{n_k} a_{n+1}^{n_1} \dots a_{2k}^{n_k} \mid n_i \in \mathbb{N}_0\}.$$

Define  $S^{(k)}$  to be  $\Phi(L^{(k)})$ . That is,

$$S^{(k)} = \{v \in \mathbb{N}_0^{(2k)} \mid v(i) = v(k+i) \ (1 \leq i \leq k)\}.$$

The following lemma gives a condition which implies a linear set is not an intersection of  $k - 1$  stratified semilinear sets. The proof is assembled primarily from the proof of [21, Lemma 4], but the result is stated differently here, because in this form it will also be useful in proving our generalisation of Liu and Weiner’s result.

**Lemma 3.4.** *Let  $S = L(\mathbf{0}; P)$  be a  $k$ -dimensional linear subset of  $\mathbb{N}_0^r$  such that  $P$  is linearly independent over  $\mathbb{Q}$ . Suppose that any subset of  $S$  which*

can be expressed as an intersection of  $k-1$  stratified linear sets with constant vector zero has dimension at most  $k-1$ . Then  $S$  is not an intersection of  $k-1$  stratified semilinear sets.

*Proof.* If  $S$  is an intersection of  $k-1$  stratified semilinear sets, then  $S$  is a finite union of intersections of  $k-1$  stratified linear sets.

Let  $L = \bigcap_{i=1}^{k-1} L_i$  be a subset of  $S$  with each  $L_i$  a stratified linear set. Let  $M = \bigcap_{i=1}^{k-1} L_i^{\mathbf{0}}$  and write  $M = L(\mathbf{0}; \mathbf{p}_1, \dots, \mathbf{p}_m)$ . By Proposition 2.8, there exists a finite subset  $C$  of  $\mathbb{N}_0^r$  such that

$$L = \bigcup_{\mathbf{c}_i \in C} L(\mathbf{c}_i; \mathbf{p}_1, \dots, \mathbf{p}_m).$$

For any  $\mathbf{c}, \mathbf{p} \in \mathbb{N}_0^r$  such that  $\mathbf{c} + n\mathbf{p} \in L$  for all  $n \in \mathbb{N}_0$ , we have  $\mathbf{p} \in L$ , since  $P$  is linearly independent over  $\mathbb{Q}$ . Thus  $M \subseteq S$ , since  $L(\mathbf{c}_1; \mathbf{p}_1, \dots, \mathbf{p}_m) \subseteq S$ .

Since  $M \subseteq S$  is an intersection of  $k-1$  stratified linear sets with constant zero,  $M$  has dimension at most  $k-1$  by the hypothesis of the lemma. Each  $L(\mathbf{c}_i; \mathbf{p}_1, \dots, \mathbf{p}_m)$  is a coset of  $M$  and thus has the same dimension as  $M$ . Thus  $L$  is a union of finitely many linear sets of dimension at most  $k-1$ . This implies that  $S$  itself is a union of finitely many linear sets of dimension at most  $k-1$ , but by Proposition 2.11, this cannot happen since  $\dim(S) = k$ .  $\square$

### 3.2.1 The new part of the proof

This subsection contains a new proof of the result which is Theorem 10 in [21], namely that  $S^{(k)}$  satisfies the hypothesis of Lemma 3.4. We break most of it up into three lemmas, which then come together to give a relatively simple proof of the proposition itself (which here is Proposition 3.8).

**Lemma 3.5.** *Let  $S = L_1 \cap \dots \cap L_k$ , where each  $L_i$  is a linear subset of  $\mathbb{N}_0^r$  with constant vector zero and periods  $P_i = \{\mathbf{p}_{i1}, \dots, \mathbf{p}_{im_i}\}$ . For each  $1 \leq i \leq k$ , let  $\mathcal{L}_i = L_i^{\mathbb{Q}}$ . If  $\dim(S) < \dim(\mathcal{L}_1 \cap \dots \cap \mathcal{L}_k)$ , then there exist  $1 \leq i \leq k$ ,  $1 \leq j \leq m_i$ , such that removing  $\mathbf{p}_{ij}$  from  $P_i$  does not change the set  $S$ .*

*Proof.* Suppose that  $\dim(S) < \dim(\mathcal{L}_1 \cap \dots \cap \mathcal{L}_k)$  and that, for all  $i$ , removing any  $\mathbf{p}_{ij}$  from  $P_i$  changes the set  $S$ . Then, for all  $i, j$ , there must exist some

$\mathbf{v}_{ij} = \alpha_{i1}^j \mathbf{p}_{i1} + \dots + \alpha_{im_i}^j \mathbf{p}_{im_i} \in S$  with  $\alpha_{ij}^j \geq 1$ .

Let  $\{\mathbf{q}_1, \dots, \mathbf{q}_s\}$  be a basis for  $\mathcal{L}_1 \cap \dots \cap \mathcal{L}_k$ . Since  $\mathbf{q}_1, \dots, \mathbf{q}_s \in \mathcal{L}_i$  for all  $i$ , we can write  $\mathbf{q}_l = \sum_{j=1}^{m_i} \beta_{ij}^l \mathbf{p}_{ij}$ , where  $\beta_{ij}^l \in \mathbb{Q}$ . Now for  $1 \leq i \leq k$ ,  $1 \leq j \leq m_i$ , let  $c_{ij} = \min\{\beta_{ij}^l \mid 1 \leq l \leq s\}$ , and let

$$\Lambda_i = \{j \mid 1 \leq j \leq m_i, c_{ij} < 0\}.$$

Then, if  $\mathbf{w}_i := \sum_{j \in \Lambda_i} -c_{ij} \mathbf{v}_{ij}$ , we have  $\mathbf{w}_i \in S$ , since  $\mathbf{v}_{ij} \in S$  and  $-c_{ij} \in \mathbb{N}$  for all  $j \in \Lambda_i$ . Each  $\mathbf{w}_i$  can thus be expressed in  $L_i$  as  $\sum_{j=1}^{m_i} \gamma_{ij} \mathbf{p}_{ij}$ , where  $\gamma_{ij} = \sum_{j' \in \Lambda_i} -c_{ij'} \alpha_{ij'}^j$ . Since  $\mathbf{w}_i$  is in  $S$ , it also has an expression

$$\mathbf{w}_i = \sum_{j=1}^{m_{i'}} \gamma_{i'j}^i \mathbf{p}_{i'j},$$

for each  $i' \neq i$  in  $\{1, \dots, k\}$ , where  $\gamma_{i'j}^i \in \mathbb{N}_0$ . For convenience, let  $\gamma_{ij}^i = \gamma_{ij}$ .

Let  $\mathbf{w} = \sum_{i=1}^k \mathbf{w}_i$ . Then  $\mathbf{w} \in S$  and, for each  $i$ , we can write

$$\mathbf{w} = \sum_{i'=1}^k \sum_{j=1}^{m_i} \gamma_{ij}^{i'} \mathbf{p}_{ij}.$$

For all  $j \in \Lambda_i$ , the coefficient of  $\mathbf{p}_{ij}$  in this expression for  $\mathbf{w}$  is

$$\sum_{i'=1}^k \gamma_{ij}^{i'} \geq \gamma_{ij}^i = \sum_{j' \in \Lambda_i} -c_{ij'} \alpha_{ij'}^j \geq -c_{ij} \alpha_{ij}^j \geq -c_{ij},$$

since  $\alpha_{ij}^j \geq 1$ . Thus we have shown that for each  $i$ , we can express  $\mathbf{w}$  in the form  $\sum_{j=1}^{m_i} a_{ij} \mathbf{p}_{ij}$ , where  $a_{ij} \geq -c_{ij}$  for all  $j \in \Lambda_i$ .

For any  $\mathbf{q}_l$  in the basis for  $\mathcal{L}_1 \cap \dots \cap \mathcal{L}_k$ , and any  $1 \leq i \leq k$ , we have

$$\mathbf{w} + \mathbf{q}_l = \sum_{j=1}^{m_i} (a_{ij} + \beta_{ij}^l) \mathbf{p}_{ij} \in L_i,$$

since  $a_{ij} + \beta_{ij}^l \geq a_{ij} + c_{ij} \geq 0$  for all  $j \in \Lambda_i$ , and  $c_{ij} \geq 0$  for  $j \notin \Lambda_i$ . Thus  $\mathbf{w} + \mathbf{q}_l \in S$  for all  $1 \leq l \leq s$ . Let  $M = \{\mathbf{w}, \mathbf{w} + \mathbf{q}_1, \dots, \mathbf{w} + \mathbf{q}_s\} \subset S$ . Then  $\mathbf{q}_1, \dots, \mathbf{q}_s$  are in the subspace of  $\mathbb{Q}^r$  generated by  $M$ , which is contained in  $S^{\mathbb{Q}}$ . Since  $\{\mathbf{q}_1, \dots, \mathbf{q}_s\}$  is a basis for  $\mathcal{L}_1 \cap \dots \cap \mathcal{L}_k$ , it is a linearly independent set over  $\mathbb{Q}$ . Thus  $S^{\mathbb{Q}}$  has at least  $s$  linearly dependent elements, contradicting  $\dim(S) = \dim(S^{\mathbb{Q}}) < \dim(\mathcal{L}_1 \cap \dots \cap \mathcal{L}_k) = s$ .  $\square$

For a stratified linear set  $L \subseteq \mathbb{N}_0^r$ , let  $\rho_L$  be the symmetric relation on  $\{1, \dots, r\}$  given by  $m\rho_L n$  if there exist non-zero  $\alpha, \beta$  with  $\alpha e_m + \beta e_n \in P$ . Define  $\sim_L$  to be the reflexive and transitive closure of  $\rho_L$ . This gives a partition  $\Pi_L$  of  $\{1, \dots, r\}$  into equivalence classes under  $\sim_L$ . Note that since  $L$  is stratified, if  $m_1 < n_1 < m_2 < n_2$ , then at most one of  $m_1\rho_L m_2$  and  $n_1\rho_L n_2$  is true. A similar property applies to  $\sim_L$ :

**Lemma 3.6.** *Let  $L \subseteq \mathbb{N}_0^r$  be a stratified linear set with constant vector zero. Then if  $m_1, n_1, m_2, n_2 \in \{1, \dots, r\}$  with  $m_1 < n_1 < m_2 < n_2$  and  $m_1 \not\sim_L n_1$ ,  $m_2 \not\sim_L n_2$ , then  $m_1 \sim_L m_2$  and  $n_1 \sim_L n_2$  cannot both occur.*

*Proof.* Suppose  $m_1 \sim_L m_2$  and  $n_1 \sim_L n_2$ . Then there exist  $i_1, \dots, i_s, j_1, \dots, j_t$  in  $\{1, \dots, r\}$  such that

$$m_1 = i_1\rho_L i_2\rho_L \dots \rho_L i_s = m_2 \quad \text{and} \quad n_1 = j_1\rho_L j_2\rho_L \dots \rho_L j_t = n_2.$$

Let  $\Lambda \in \Pi_L$  such that  $m_1, m_2 \in \Lambda$ . Then since  $m_1 < n_1 < m_2 < n_2$  and  $n_1, n_2 \notin \Lambda$ , there must exist  $k$  such that either  $m_1 < j_k < m_2 < j_{k+1}$ , or  $j_{k+1} < m_1 < j_k < m_2$ . Since  $j_k\rho_L j_{k+1}$ , this forces  $i_l$  to lie between  $j_k$  and  $j_{k+1}$  for all  $1 \leq l \leq s$ . But either  $m_1 (= i_1)$  or  $m_2 (= i_s)$  does not lie between  $j_k$  and  $j_{k+1}$ , thus we have a contradiction.  $\square$

The following result gives a relationship between  $\Pi_L$  and the orthogonal complement of  $L^\mathbb{Q}$ .

**Lemma 3.7.** *Let  $L \subseteq \mathbb{N}_0^r$  be a stratified linear set, with  $\Pi_L = \{\Lambda_1, \dots, \Lambda_t\}$ , and let  $\mathcal{L} = L^\mathbb{Q}$ . Then  $\mathcal{L}^\perp$  has a basis of the form  $\{\mathbf{x}_i = \sum_{j \in \Lambda_i} \gamma_j e_j \mid i \in M\}$ , where  $M \subseteq \{1, \dots, t\}$ . In particular,  $\dim(\mathcal{L}^\perp) = |M| \leq t$ .*

*Proof.* Let  $M$  be the set of all  $i \in \{1, \dots, t\}$  such that  $\mathbf{x}(j) \neq 0$  for some  $\mathbf{x} \in \mathcal{L}^\perp$  and  $j \in \Lambda_i$ . For each  $i \in M$ , fix some non-zero  $\mathbf{x}^{(i)} \in \mathcal{L}^\perp$  with  $\mathbf{x}^{(i)}(j) \neq 0$  for some  $j \in \Lambda_i$ . We can write  $\mathbf{x}^{(i)} = \sum_{j=1}^r \gamma_{ij} e_j = \sum_{s=1}^t \mathbf{x}_s^{(i)}$ , where  $\mathbf{x}_s^{(i)} = \sum_{j \in \Lambda_s} \gamma_{ij} e_j$ , since  $\{1, \dots, r\}$  is the disjoint union of  $\Lambda_1, \dots, \Lambda_t$ . For  $i \in M$ , let  $\mathbf{x}_i = \mathbf{x}_i^{(i)}$ . Then  $\{\mathbf{x}_i \mid i \in M\}$  is a linearly independent set, since  $\mathbf{x}_i \neq 0$  by the choice of  $\mathbf{x}^{(i)}$ , and  $\mathbf{x}_i(j) = 0$  for all  $j \notin \Lambda_i$ .

Let  $P$  be the set of periods of  $L$ , and for  $1 \leq i \leq t$ , let

$$P_i = \{\alpha_m e_m + \alpha_n e_n \in P \mid m, n \in \Lambda_i\},$$



where one of  $\alpha_m$  or  $\alpha_n$  may be zero. Then  $\{P_1, \dots, P_t\}$  is a partition of  $P$ . Now if  $\mathbf{p} \in P_i$ , then  $\mathbf{p} \cdot \mathbf{x}_{i'}^{(i)} = 0$  for all  $i' \neq i$ , since  $\mathbf{x}_{i'}^{(i)}(j) = 0$  for all  $j \in \Lambda_i$ . Thus  $\mathbf{p} \cdot \mathbf{x}^{(i)} = \mathbf{p} \cdot (\mathbf{x}_1^{(i)} + \dots + \mathbf{x}_t^{(i)}) = \mathbf{p} \cdot \mathbf{x}_i^{(i)} = \mathbf{p} \cdot \mathbf{x}_i$ . But  $\mathbf{x}^{(i)} \in \mathcal{L}^\perp$ , so  $\mathbf{p} \cdot \mathbf{x}_i = 0$ . Since also  $\mathbf{p} \cdot \mathbf{x}_i = 0$  for all  $\mathbf{p} \in P_{i'}$  with  $i' \neq i$ , we have  $\mathbf{x}_i \in \mathcal{L}^\perp$ , for all  $i \in M$ .

It remains to show that  $\{\mathbf{x}_i \mid i \in M\}$  spans  $\mathcal{L}^\perp$ . Recall that  $\mathbf{x}_i = \sum_{j \in \Lambda_i} \gamma_{ij} e_j$ . First we show that  $\gamma_{ij} \neq 0$  for all  $i \in M, j \in \Lambda_i$ . For  $i \in M$ , certainly  $\gamma_{im} \neq 0$  for some  $m \in \Lambda_i$ , since  $\mathbf{x}_i \neq \mathbf{0}$ . For any  $n \in \Lambda_i$  there exist  $m_1, \dots, m_l \in \Lambda_i$  such that  $m = m_1 \rho_L m_2 \rho_L \dots \rho_L m_l = n$ , which implies the existence of periods  $\alpha_{m_1} e_{m_1} + \alpha_{m_2} e_{m_2}, \dots, \alpha_{m_{l-1}} e_{m_{l-1}} + \alpha_{m_l} e_{m_l} \in P_i$  with non-zero  $\alpha_{m_j}$  for all  $1 \leq j \leq l$ . Now

$$\mathbf{x}_i \cdot (\alpha_{m_j} e_{m_j} + \alpha_{m_{j+1}} e_{m_{j+1}}) = \gamma_{im_j} \alpha_{m_j} + \gamma_{im_{j+1}} \alpha_{m_{j+1}} = 0$$

for all  $1 \leq j \leq l-1$ , since  $\mathbf{x}_i \in \mathcal{L}^\perp$ . Thus  $\gamma_{im_{j+1}} = -\gamma_{im_j} \frac{\alpha_{im_j}}{\alpha_{im_{j+1}}}$  and so by induction  $\gamma_{in} = \gamma_{im_l} \neq 0$ , since  $\gamma_{im} = \gamma_{i1} \neq 0$ . Moreover, for all  $n \in \Lambda_i$ , the coefficient  $\gamma_{in}$  is uniquely determined by  $\gamma_{im}$ . (If two different paths between  $m$  and  $n$  gave different values for  $\gamma_{in}$ , then our non-zero  $\mathbf{x}_i \in \mathcal{L}^\perp$  could not exist.)

Finally, let  $\mathbf{y} \in \mathcal{L}^\perp$  and write  $\mathbf{y} = \sum_{j=1}^r c_j e_j = \sum_{i=1}^t \mathbf{y}_i$ , where  $\mathbf{y}_i = \sum_{j \in \Lambda_i} c_j e_j$ . If  $\mathbf{y}_i \neq \mathbf{0}$ , then choose  $j \in \Lambda_i$  with  $c_j \neq 0$ . Since  $\gamma_{ij} \neq 0$ , we can write  $c_j = q \gamma_{ij}$ , where  $q \in \mathbb{Q}$ . By exactly the same argument as we used for  $\mathbf{x}_i$ , we can conclude that  $\mathbf{p} \cdot \mathbf{y}_i = 0$  for all  $\mathbf{p} \in P$ . Now for any  $\alpha_j e_j + \alpha_{j'} e_{j'} \in P_i$ , we have  $\mathbf{y}_i \cdot (\alpha_j e_j + \alpha_{j'} e_{j'}) = c_j \alpha_j + c_{j'} \alpha_{j'}$ , thus  $c_{j'} = -c_j \frac{\alpha_j}{\alpha_{j'}}$ . But also  $\gamma_{ij'} = -\gamma_{ij} \frac{\alpha_j}{\alpha_{j'}}$ . Thus  $c_{j'} = -q \gamma_{ij} \frac{\alpha_j}{\alpha_{j'}} = q \gamma_{ij'}$ , and we can extend this to show that  $c_n = q \gamma_{in}$  for all  $n \in \Lambda_i$ , thus  $\mathbf{y}_i = q \mathbf{x}_i$ . Since this applies to all  $i \in M$  with  $\mathbf{y}_i \neq \mathbf{0}$ , we can conclude that  $\mathbf{y}$  is a linear combination of the elements of  $\{\mathbf{x}_i \mid i \in M\}$ , and thus this set spans  $\mathcal{L}^\perp$ .  $\square$

We are now ready to prove Theorem 10 of [21].

**Proposition 3.8.** *For  $1 \leq i \leq k-1$ , let  $L_i$  be a stratified linear set with constant vector zero, and let  $L_1 \cap \dots \cap L_{k-1} = S \subseteq S^{(k)}$ . Then  $S$  is a linear set of dimension at most  $k-1$ .*

*Proof.*  $S$  is a linear set with constant vector zero by Proposition 2.8. Let  $\mathcal{L}_i = L_i^\mathbb{Q}$  for all  $1 \leq i \leq k-1$ , and let  $\mathcal{S} = \mathcal{L}_1 \cap \dots \cap \mathcal{L}_{k-1}$ . By Lemma 3.5, we can assume that  $\dim(\mathcal{S}) = \dim(S)$ . Since  $S \subseteq \mathcal{S}$ , this implies that any

maximal linearly independent subset of the periods of  $S$  is a basis for  $\mathcal{S}$ . Thus, since  $\mathbf{v}(i) = \mathbf{v}(k+i)$  for all  $\mathbf{v} \in S$ , we also have  $\mathbf{v}(i) = \mathbf{v}(k+i)$  for all  $\mathbf{v} \in \mathcal{S}$ . For all  $1 \leq i \leq k$ , we have  $e_i - e_{k+i} \in \mathcal{S}^\perp$ , since  $\mathbf{v} \cdot (e_i - e_{k+i}) = \mathbf{v}(i) - \mathbf{v}(k+i) = 0$  for all  $\mathbf{v} \in \mathcal{S}$ .

Assume  $\{e_i - e_{k+i} \mid 1 \leq i \leq k\}$  spans  $\mathcal{S}^\perp$ , since otherwise  $\dim(\mathcal{S}^\perp) \geq k+1$  and thus  $\dim(\mathcal{S}) \leq 2k - (k+1) = k-1$ .

If  $\mathcal{L}_i^\perp \neq \{\mathbf{0}\}$ , let  $\Pi_{L_i} = \{\Lambda_1, \dots, \Lambda_t\}$ . Then, by Lemma 3.7,  $\mathcal{L}_i^\perp$  has a basis of the form  $\{\mathbf{x}_s \mid s \in M\}$ , where  $M \subseteq \{1, \dots, t\}$  and  $\mathbf{x}_s = \sum_{j \in \Lambda_s} \gamma_j e_j$ . If  $s \in M$ , then since  $\mathbf{x}_s \in \mathcal{L}_i^\perp \subseteq \mathcal{L}^\perp$ , we can write

$$\mathbf{x}_s = \sum_{j \in \Gamma_s} \gamma_j (e_j - e_{k+j}),$$

where  $\Gamma_s = \Lambda_s \cap \{1, \dots, k\}$ . Certainly some  $\gamma_j$  must be non-zero, implying  $j, (k+j) \in \Lambda_s$ . Thus if  $s, s' \in M$ , then we would have some  $j, (k+j) \in \Lambda_s$ ,  $l, (k+l) \in \Lambda_{s'}$ . But either  $j < l < (k+j) < (k+l)$  or  $l < j < (k+l) < (k+j)$ , thus this would contradict Lemma 3.6. Therefore at most one  $s \in M$ , and so  $\dim(\mathcal{L}_i^\perp) \leq 1$ . This holds for all  $1 \leq i \leq k-1$ .

But if each  $\mathcal{L}_i^\perp$  is at most one dimensional, then since  $\mathcal{S}^\perp = \mathcal{L}_1^\perp + \dots + \mathcal{L}_{k-1}^\perp$ ,  $\dim(\mathcal{S}^\perp)$  cannot exceed  $k-1$ , contradicting the fact that  $e_j - e_{k+j} \in \mathcal{S}^\perp$  for all  $1 \leq j \leq k$ . Thus our assumption that  $\{e_j - e_{k+j} \mid 1 \leq j \leq k\}$  spans  $\mathcal{S}^\perp$  was false, and so in fact  $\dim(\mathcal{S}) \leq k-1$ .  $\square$

### 3.2.2 The rest of the proof

**Theorem 3.9.** [21, Theorem 8] *The language  $L^{(k)}$  is  $k$ - $\mathcal{CF}$ , but not  $(k-1)$ - $\mathcal{CF}$ . Thus, for all  $k \geq 2$ , the class of  $k$ - $\mathcal{CF}$  languages properly contains the class of  $(k-1)$ - $\mathcal{CF}$  languages.*

*Proof.* By Corollary 2.7, it suffices to show that  $S^{(k)}$  is an intersection of  $k$  but not  $k-1$  stratified semilinear sets. For  $1 \leq i \leq k$ , define

$$S_i = \text{span} \{e_i + e_{k+i}, e_j \mid 1 \leq j \leq 2k, j \notin \{i, k+i\}\}.$$

Then each  $S_i$  is a stratified linear set and  $S^{(k)} = \bigcap_{i=1}^k S_i$ . Also,  $S^{(k)}$  has constant vector zero and dimension  $k$ , since  $\{e_i + e_{k+i} \mid 1 \leq i \leq k\}$  is a linearly independent subset which spans  $S^{(k)}$ . Hence, by Proposition 3.8,

$S^{(k)}$  satisfies the hypothesis of Lemma 3.4, so cannot be expressed as an intersection of  $k - 1$  stratified semilinear sets.  $\square$

### 3.3 The languages $L^{(n,k)}$

We can extend Theorem 3.9 to a larger, but very similar, class of languages. The extended result will be used to prove that certain groups, for example the restricted standard wreath products  $C_p \wr \mathbb{Z}$  (for any  $p > 1$ ), are not poly- $\mathcal{CF}$ .

For each  $n, k \in \mathbb{N}$ , let  $a_1, a_2, \dots, a_{2nk}$  be  $2nk$  distinct symbols and define

$$L^{(n,k)} = \{a_1^{m_1} a_2^{m_2} \dots a_{2nk}^{m_{2nk}} \mid \begin{array}{l} m_i \in \mathbb{N}_0, m_i = m_{nk+i} \ (1 \leq i \leq nk), \\ m_{nj+1} = m_{nj+l} \ (0 \leq j \leq k-1, 2 \leq l \leq n) \end{array}\}.$$

For example,  $L^{(2,2)} = \{a_1^m a_2^m a_3^n a_4^n a_5^m a_6^m a_7^n a_8^n \mid m, n \in \mathbb{N}_0\}$ . Define  $S^{(n,k)}$  to be  $\Phi(L^{(n,k)})$ . Then

$$S^{(n,k)} = \{\mathbf{v} \in \mathbb{N}_0^{2nk} \mid \begin{array}{l} \mathbf{v}(i) = \mathbf{v}(i + nk) \ (1 \leq i \leq nk), \\ \mathbf{v}(nj + 1) = \mathbf{v}(nj + l) \ (0 \leq j \leq k-1, 2 \leq l \leq n) \end{array}\}.$$

These sets are like  $S^{(k)}$ , except with each entry being repeated  $n$  times. Thus  $S^{(1,k)}$  is just  $S^{(k)}$ . For any  $n \in \mathbb{N}$ , the set  $S^{(n,k)}$  has dimension  $k$ , so it is not surprising that the following result does not depend on  $n$ .

**Proposition 3.10.** *For  $1 \leq i \leq k - 1$ , let  $L_i$  be a stratified linear set with constant vector zero, and let  $L_1 \cap \dots \cap L_{k-1} = S \subseteq S^{(n,k)}$ . Then  $S$  is a linear set of dimension at most  $k - 1$ .*

*Proof.* The proof follows the idea of the proof of Proposition 3.8, but is a good deal more complicated.

$S$  is a linear set with constant vector zero by Proposition 2.8. Let  $\mathcal{L}_i = L_i^{\mathbb{Q}}$  for  $1 \leq i \leq k - 1$ , and let  $\mathcal{S} = \mathcal{L}_1 \cap \dots \cap \mathcal{L}_{k-1}$ . By Lemma 3.5, we can assume that  $\dim(\mathcal{S}) = \dim(S)$ . Since  $S \subseteq \mathcal{S}$ , this implies that any maximal linearly independent subset of the periods of  $S$  is a basis for  $\mathcal{S}$ .

Thus since  $\mathbf{v}(i) = \mathbf{v}(nk+i)$  for all  $\mathbf{v} \in S$ , we also have  $\mathbf{v}(i) = \mathbf{v}(nk+i)$  for all  $\mathbf{v} \in \mathcal{S}$ ,  $1 \leq i \leq nk$ . Moreover, for all  $\mathbf{v} \in \mathcal{S}$  we have  $\mathbf{v}(nj+l) = \mathbf{v}(nj+l+1)$  for all  $0 \leq j \leq k-1$ ,  $1 \leq l \leq n-1$ .

For all  $1 \leq i \leq nk$ , we have  $e_i - e_{nk+i} \in \mathcal{S}^\perp$ , since, for all  $\mathbf{v} \in \mathcal{S}$ ,

$$\mathbf{v} \cdot (e_i - e_{nk+i}) = \mathbf{v}(i) - \mathbf{v}(nk+i) = 0.$$

Similarly,  $e_{nj+l} - e_{nj+l+1} \in \mathcal{S}^\perp$  for all  $0 \leq j \leq k-1$  and  $1 \leq l \leq n-1$ . Thus we know of  $nk + (n-1)k = (2n-1)k$  linearly independent elements of  $\mathcal{S}^\perp$ .

Assume that these  $(2n-1)k$  elements form a basis of  $\mathcal{S}^\perp$ , since otherwise

$$\dim(S) = \dim(\mathcal{S}) < 2nk - (2n-1)k = k,$$

as we require. We will now derive a contradiction, using the fact that  $\mathcal{S}^\perp = \mathcal{L}_1^\perp + \dots + \mathcal{L}_{k-1}^\perp$ .

For  $0 \leq j \leq k-1$  and  $\epsilon \in \{0, 1\}$ , define

$$\Delta_j^\epsilon = \{n(\epsilon k + j) + l \mid 1 \leq l \leq n\}$$

and  $\Delta_j = \Delta_j^0 \cup \Delta_j^1$ . Let  $\mathcal{S}_j$  be the image of the projection of  $\mathcal{S}^\perp$  onto the coordinates in  $\Delta_j$ . Since every vector in the basis of  $\mathcal{S}^\perp$  above is contained in some  $\mathcal{S}_j$ , and the  $\Delta_j$  are disjoint,  $\mathcal{S}^\perp$  is the direct sum of  $\mathcal{S}_0, \dots, \mathcal{S}_{k-1}$ .

Call  $\mathbf{x} \in \mathcal{S}^\perp$  a *j-bridge* if there exist  $l \in \Delta_j^0$  and  $l' \in \Delta_j^1$  such that  $\mathbf{x}(l)$  and  $\mathbf{x}(l')$  are both non-zero. By extension, for  $\Gamma \subseteq \{0, \dots, k-1\}$ , call  $\mathbf{x}$  a  *$\Gamma$ -bridge* if  $\mathbf{x}$  is a *j-bridge* for all  $j \in \Gamma$ .

For  $0 \leq j \leq k-1$ , let  $\Omega_j$  be the  $2(n-1)$ -dimensional subspace of  $\mathcal{S}_j^\perp$  generated by

$$\{e_{n(\epsilon k + j) + l} - e_{n(\epsilon k + j) + l + 1} \mid \epsilon \in \{0, 1\}, 1 \leq l \leq n-1\}$$

and let  $\Omega = \Omega_0 + \dots + \Omega_{k-1}$ .

Suppose that  $\mathbf{x}$  is not a *j-bridge* for any  $j$ . We will show that  $\mathbf{x}$  must be in  $\Omega$ . Write  $\mathbf{x} = \sum_{j=0}^{k-1} \mathbf{y}_j$ , where  $\mathbf{y}_j \in \mathcal{S}_j$ . Then no  $\mathbf{y}_j$  is a *j-bridge*. For any  $j \in \{0, \dots, k-1\}$ , the non-zero coordinates of  $\mathbf{y}_j$  are either all in  $\Delta_j^0$  or all in  $\Delta_j^1$ , since  $\mathbf{y}_j$  is in  $\mathcal{S}_j$  and is not a *j-bridge*. For any  $\mathbf{v} \in \mathcal{S}^\perp$ , the sum of the entries of  $\mathbf{v}$  is zero, as can be seen by considering the basis vectors of  $\mathcal{S}^\perp$ . Thus the subspace of  $\mathcal{S}^\perp$  consisting of vectors whose non-zero coordinates all lie in  $\Delta_j^\epsilon$  is spanned by  $\{e_{n(\epsilon k + j) + l} - e_{n(\epsilon k + j) + l + 1} \mid 1 \leq l \leq n-1\} \subseteq \Omega_j$ , for  $\epsilon \in \{0, 1\}$ . Hence  $\mathbf{y}_j \in \Omega_j$ , and since this applies for all  $0 \leq j \leq k-1$ , we conclude that  $\mathbf{x} = \sum_{j=0}^{k-1} \mathbf{y}_j \in \Omega$ .

If  $\mathcal{L}_i^\perp \neq \{\mathbf{0}\}$ , let  $\Pi_{L_i} = \{\Lambda_1, \dots, \Lambda_t\}$ . Then by Lemma 3.7,  $\mathcal{L}_i^\perp$  has a basis of the form  $B_i = \{\mathbf{x}_s \mid s \in M\}$ , where  $M \subseteq \{1, \dots, t\}$  and  $\mathbf{x}_s = \sum_{j \in \Lambda_s} \gamma_{sj} e_j$ . Note that if  $\mathbf{x}_s$  is a  $j$ -bridge and  $s' \neq s$ ,  $j' \neq j$ , then  $\mathbf{x}_{s'}$  cannot be a  $j'$ -bridge, since this would imply the existence of  $l_1, l_2, l'_1, l'_2 \in \{1, \dots, n\}$  such that

$$nj + l_1, n(k + j) + l_2 \in \Lambda_s, \quad nj' + l'_1, n(k + j') + l'_2 \in \Lambda_{s'},$$

contradicting Lemma 3.6.

If  $B_i$  contains no  $\Gamma$ -bridges for any non-empty  $\Gamma$ , then every  $\mathbf{x}_s \in B_i$  is in  $\Omega$ , hence  $\mathcal{L}_i^\perp \subseteq \Omega$ . If the largest  $\Gamma$  such that  $\mathbf{x}_s$  is a  $\Gamma$ -bridge is a singleton  $\{j\}$ , then  $B_i$  may possibly contain other  $j$ -bridges; but, as already observed,  $B_i$  contains no  $j'$ -bridges for  $j' \neq j$ . If  $\Gamma$  has at least two elements and  $\mathbf{x}_s$  is a  $\Gamma$ -bridge, then  $B_i$  contains no other  $\Gamma'$ -bridges, even for  $\Gamma' = \Gamma$ , since this would again imply a situation contradicting Lemma 3.6.

Thus there is at most one  $\Gamma \subseteq \{0, \dots, k-1\}$  such that  $B_i$  contains one or more  $\Gamma$ -bridges. If such  $\Gamma$  exists, call it  $\Gamma_i$ .

For each  $i$ , we have  $\mathcal{L}_i^\perp = \mathcal{M}_i + \mathcal{N}_i$ , where  $\mathcal{M}_i$  is the subspace generated by the  $\Gamma_i$ -bridge(s) and  $\mathcal{N}_i$  is the subspace generated by the remaining elements of  $B_i$ .

Now consider

$$\mathcal{S}^\perp = \mathcal{L}_1^\perp + \dots + \mathcal{L}_{k-1}^\perp = \mathcal{M}_1 + \dots + \mathcal{M}_{k-1} + \mathcal{N}_1 + \dots + \mathcal{N}_{k-1}.$$

Since the  $\mathcal{N}_i$  are generated by elements which are not  $\Gamma$ -bridges for any non-empty  $\Gamma$ , they are all subspaces of  $\Omega$ . Thus  $\mathcal{S}^\perp \subseteq \mathcal{M}_1 + \dots + \mathcal{M}_{k-1} + \Omega$ .

If  $\Gamma_i$  contains at least two elements, then  $B_i$  has a single  $\Gamma_i$ -bridge, so  $\mathcal{M}_i$  has dimension one. If  $\Gamma_i = \{j\}$ , then even though  $\mathcal{M}_i$  can have dimension up to  $n$ ,  $\Omega_j + \mathcal{M}_i$  has to be contained in  $\mathcal{S}_j$ , so can have dimension at most  $2n - 1$ , which is one more than the dimension of  $\Omega_j$ . Thus each  $\mathcal{M}_i$  contributes at most one extra dimension to the set  $\Omega + \mathcal{M}_1 + \dots + \mathcal{M}_{k-1}$ , and so

$$\begin{aligned} \dim(\mathcal{S}^\perp) &\leq \dim(\Omega + \mathcal{M}_1 + \dots + \mathcal{M}_{k-1}) \\ &\leq 2k(n-1) + k-1 = (2n-1)k-1, \end{aligned}$$

giving a contradiction. Thus our assumption that  $\mathcal{S}^\perp$  was spanned by  $(2n-1)k$  elements is incorrect, and so

$$\dim(\mathcal{S}) = \dim(\mathcal{S}) \leq 2nk - ((2n-1)k+1) = k-1. \quad \square$$

**Corollary 3.11.** *A  $k$ -dimensional linear subset of  $S^{(n,k)}$  cannot be expressed as an intersection of  $k - 1$  stratified semilinear sets.*

*Proof.* Suppose  $L \subseteq S^{(n,k)}$  is  $k$ -dimensional and can be expressed as an intersection of  $k - 1$  stratified semilinear sets. Then we can write  $L = S_1 \cup \dots \cup S_l$ , where each  $S_i$  is an intersection of  $k - 1$  stratified linear sets. By Proposition 2.8, there exist finite subsets  $C_i$  and  $P_i$  of  $\mathbb{N}_0^{2nk}$  such that  $S_i = L(C_i; P_i)$  for  $1 \leq i \leq l$ . By Proposition 2.11, there must exist  $1 \leq i \leq l$  and  $\mathbf{c} \in C_i$  such that  $L(\mathbf{c}; P_i)$  has dimension  $k$ , and hence  $L(\mathbf{0}; P_i)$  has dimension  $k$ . Writing  $S_i = \cap_{i=1}^{k-1} N_i$ , where each  $N_i$  is a stratified linear set, from Proposition 2.8 we have  $L(\mathbf{0}; P_i) = \cap_{i=1}^{k-1} N_i^{\mathbf{0}}$ . But  $L(\mathbf{0}; P_i)$  is a  $k$ -dimensional linear subset of  $S^{(n,k)}$  with constant zero, while each  $N_i^{\mathbf{0}}$  is a stratified linear set, contradicting Proposition 3.10.  $\square$

**Theorem 3.12.** *For any  $k, n \in \mathbb{N}$ , the set  $S^{(n,k)}$  is not an intersection of  $k - 1$  stratified semilinear sets, and so the language  $L^{(n,k)}$  is not  $(k - 1)$ - $\mathcal{CF}$ .*

*Proof.* Recall from the proof of Proposition 3.10 the notation

$$\Delta_j = \{nj + l \mid 1 \leq l \leq n\} \cup \{n(k + j) + l \mid 1 \leq l \leq n\}.$$

For  $0 \leq j \leq k - 1$ , let  $\mathbf{u}_j = \sum_{i \in \Delta_j} e_i$ . Then  $\{\mathbf{u}_j \mid 0 \leq j \leq k - 1\}$  is a linearly independent set which spans  $S^{(n,k)}$ , so  $S^{(n,k)}$  is  $k$ -dimensional. Since  $S^{(n,k)}$  has constant vector zero, it follows from Lemma 3.4 and Proposition 3.10 that  $S^{(n,k)}$  cannot be an intersection of  $k - 1$  stratified semilinear sets and thus  $L^{(n,k)}$  cannot be a  $(k - 1)$ - $\mathcal{CF}$  language.  $\square$

## 4 Poly- $\mathcal{CF}$ groups

We begin with a simple observation, followed by our main conjecture.

**Observation 4.1.** *The class of poly- $\mathcal{CF}$  groups is closed under taking finite direct products. The direct product of a  $k_1$ - $\mathcal{CF}$  group and a  $k_2$ - $\mathcal{CF}$  group is  $(k_1 + k_2)$ - $\mathcal{CF}$ .*

*Proof.* It suffices to show that the direct product of two poly- $\mathcal{CF}$  groups is poly- $\mathcal{CF}$ . Let  $G_i$  be a  $k_i$ - $\mathcal{CF}$  group for  $i = 1, 2$ . Let  $A_{i1}, \dots, A_{ik_i}$  be pushdown automata with input alphabet  $X_i$  such that a word is in  $W(G_i, X_i)$

if and only if it is accepted by all  $A_{ij}$ . We may assume that  $X_1$  and  $X_2$  are disjoint. Now modify the automata  $A_{ij}$  so that their input alphabet is  $X = X_1 \cup X_2$ , but each  $A_{1j}$  ignores the symbols in  $X_2$  and  $A_{2j}$  ignores the symbols in  $X_1$ . Let  $h_1 : X \rightarrow X_1$  be the homomorphism sending every symbol in  $X_2$  to the empty word, and define  $h_2$  similarly. Then a word  $w$  in  $(X \cup X^{-1})^*$  is accepted by all of the modified automata  $A_{ij}$  if and only if  $h_i(w) \in W(G_i, X_i)$  for  $i = 1, 2$ . Thus the intersection of the languages accepted by all the  $A_{ij}$  is precisely  $W(G_1 \times G_2, X)$ , and hence  $G_1 \times G_2$  is  $(k_1 + k_2)\text{-}\mathcal{CF}$ .  $\square$

Since finitely generated free groups are context-free, this implies that a direct product of  $k$  finitely generated free groups is  $k\text{-}\mathcal{CF}$ . Since the  $k\text{-}\mathcal{CF}$  groups are closed under taking finite index overgroups and finitely generated subgroups, any finitely generated subgroup of a direct product of  $k$  free groups, and any finite index overgroup of such a group, is  $k\text{-}\mathcal{CF}$ . These are the only known  $k\text{-}\mathcal{CF}$  groups, and we conjecture that they are the only ones.

**Conjecture 4.2.** *Let  $G$  be a finitely generated group. Then  $G$  is poly- $\mathcal{CF}$  if and only if  $G$  is virtually a finitely generated subgroup of a direct product of free groups.*

This would generalise both Muller and Schupp's result on context-free groups [23, 24, 4] and the theorem of Holt, Owens and Thomas [15], which says that the word problem of a finitely generated group is an intersection of finitely many one-counter languages if and only if the group is virtually abelian. A *one-counter language* is a language recognised by a pushdown automaton with only one stack symbol.

Note that the truth of Conjecture 4.2 would imply that if  $G$  is poly- $\mathcal{CF}$ , then  $W(G)$  is an intersection of finitely many deterministic context-free languages, and hence  $\text{co}W(G)$  is context-free, since the deterministic context-free languages are closed under complementation and the context-free languages are closed under union.

The rest of this section is devoted to proving certain classes of groups to be *not* poly- $\mathcal{CF}$ .

#### 4.1 Some groups which are not poly- $\mathcal{CF}$

Holt, Rees, Röver and Thomas proved that a finitely generated nilpotent group or polycyclic group is  $\text{co}\mathcal{CF}$  if and only if it is virtually abelian [11, Theorems 12 and 16], and that the Baumslag-Solitar group  $\text{BS}(m, n)$  is not  $\text{co}\mathcal{CF}$  if  $m \neq \pm n$  [11, Theorem 13]. These theorems are all proved using [11, Proposition 14], which, as we have mentioned, has a strictly weaker hypothesis than Proposition 3.2; so, with no further effort, we can obtain analogous results for poly- $\mathcal{CF}$  groups, using Corollary 3.3.

**Proposition 4.3.** *Let  $G$  be a polycyclic group or a finitely generated nilpotent group. Then  $G$  is poly- $\mathcal{CF}$  if and only if it is virtually abelian.*

*Proof.* If  $G$  is not virtually abelian, then the proofs of Theorems 12 (for  $G$  nilpotent) and 16 (for  $G$  polycyclic) in [11] show that there exists a regular language  $R$  such that  $\phi(W(G) \cap R)$  satisfies the hypothesis of Proposition 3.2, and hence  $G$  is neither  $\text{co}\mathcal{CF}$  nor poly- $\mathcal{CF}$  by Corollary 3.3.  $\square$

The result for nilpotent groups was actually already obtained by Holt, Owens and Thomas in [15], using what is essentially a special case of Proposition 3.2.

The statement of Theorem 13 in [11] is incorrect. It is claimed that  $\text{BS}(m, n)$  is  $\text{co}\mathcal{CF}$  if and only if it is virtually abelian, based on the supposition that  $\text{BS}(m, n)$  is virtually abelian if  $m = \pm n$ . We now show that if  $m = \pm n$ , then  $\text{BS}(m, n)$  is both  $\text{co}\mathcal{CF}$  and poly- $\mathcal{CF}$ .

**Proposition 4.4.** *For  $m \in \mathbb{Z} \setminus \{0\}$ , the Baumslag-Solitar group  $\text{BS}(m, \pm m)$  is virtually a direct product of two free groups and is thus both  $\text{co}\mathcal{CF}$  and 2- $\mathcal{CF}$ .*

*Proof.* First let  $G = \text{BS}(m, m) = \langle x, y \mid y^{-1}x^my = x^m \rangle$ . Then  $x^m \in Z(G)$  and

$$G/\langle x^m \rangle = \langle x, y \mid x^m \rangle = C_m * \mathbb{Z}.$$

Let  $H/\langle x^m \rangle$  be the normal closure in  $G/\langle x^m \rangle$  of  $\langle y \rangle$ . Then

$$|G/\langle x^m \rangle : H/\langle x^m \rangle| = m$$

and hence  $|G : H| = m$ . Since  $H/\langle x^m \rangle$  does not intersect any conjugate of  $C_m$ , by the Kurosh Subgroup Theorem (see for example [22, III.3.6]),



$H/\langle x^m \rangle$  is the free product of a free group with conjugates of  $\mathbb{Z}$ , and is thus free. Since  $x^m \in Z(G)$ , we have  $H \cong H/\langle x^m \rangle \times \langle x^m \rangle$ . Thus  $G$  is virtually a direct product of two free groups.

Now let  $G = \text{BS}(m, -m) = \langle x, y \mid y^{-1}x^my = x^{-m} \rangle$ . Let  $K$  be the normal closure in  $G$  of  $\langle x, y^2 \rangle$ , which has index 2 in  $G$ . Setting  $a = x$ ,  $b = y^{-1}x^{-1}y$  and  $c = y^2$  gives

$$K = \langle a, b, c \mid a^m = b^m, [a^m, c] \rangle,$$

with  $a^m \in Z(K)$ . Now take

$$H := K/\langle a^m \rangle = \langle a, b, c \mid a^m = b^m = 1 \rangle = C_m * C_m * \mathbb{Z}.$$

Let  $\phi$  be the homomorphism from  $H$  to  $C_m \times C_m$  given by mapping  $a$  onto a generator of the first  $C_m$  and  $b$  onto a generator of the second  $C_m$ , and  $c$  onto the identity. Then the intersection of  $\ker \phi$  with every conjugate of  $\langle a \rangle$  and  $\langle b \rangle$  is trivial. Thus  $\ker \phi$  is free, again by the Kurosh Subgroup Theorem. Also,  $|H : \ker \phi| = |C_m \times C_m| = m^2$ . Let  $K_1$  be the preimage of  $\ker \phi$  in  $K$ . Since  $\ker \phi$  is free and  $\langle a^m \rangle \in Z(H)$ ,  $K_1$  is isomorphic to  $\ker \phi \times \langle a^m \rangle$ . Also,  $K_1$  has finite index in  $K$ , and hence also in  $G$ , since  $\ker \phi$  has finite index in  $H = K/\langle a^m \rangle$ . Thus  $G$  is virtually a direct product of two free groups.

Hence  $G$  is 2- $\mathcal{CF}$  by Observation 4.1, and  $\text{co}\mathcal{CF}$  by the fact that the  $\text{co}\mathcal{CF}$  groups are closed under taking finite direct products [11, Proposition 6].  $\square$

We can now determine which Baumslag-Solitar groups are poly- $\mathcal{CF}$ .

**Proposition 4.5.** *The Baumslag-Solitar group  $\text{BS}(m, n)$  is poly- $\mathcal{CF}$  or  $\text{co}\mathcal{CF}$  if and only if  $m = \pm n$ .*

*Proof.* The proof of Theorem 13 in [11] shows that if  $G = \text{BS}(m, n)$  with  $m \neq \pm n$ , then  $W(G)$  can be intersected with a regular language to give a sublanguage satisfying the hypothesis of Proposition 3.2, and so  $W(G)$  is neither  $\text{co}\mathcal{CF}$  nor poly- $\mathcal{CF}$  by Corollary 3.3.  $\square$

## 4.2 Free abelian groups and wreath products

The obvious application of Proposition 3.9 to word problems of groups is to the free abelian groups.

**Lemma 4.6.** *A free abelian group of rank  $k$  is  $k$ - $\mathcal{CF}$  but not  $(k - 1)$ - $\mathcal{CF}$ .*

*Proof.* The group  $\mathbb{Z}^k$  is a direct product of  $k$  free groups, and is hence  $k\text{-}\mathcal{CF}$ . Let  $\{x_1, \dots, x_k\}$  be a generating set for  $\mathbb{Z}^k$  and let  $X_i$  denote the inverse of  $x_i$ . Consider  $L = W(\mathbb{Z}^k) \cap (x_1^* \dots x_k^* X_1^* \dots X_k^*)$ . This is precisely the language  $L^{(k)} = \{x_1^{n_1} \dots x_k^{n_k} X_1^{n_1} \dots X_k^{n_k} \mid n_i \in \mathbb{N}_0\}$  defined in Section 3.2. Thus, by Proposition 3.9,  $L$  is not  $(k-1)\text{-}\mathcal{CF}$ . Since  $L$  is the intersection of  $W(\mathbb{Z}^k)$  with a regular language, this implies that  $\mathbb{Z}^k$  is not  $(k-1)\text{-}\mathcal{CF}$ .  $\square$

The class of  $\text{co}\mathcal{CF}$  groups is closed under taking restricted standard wreath products with context-free top group [11, Theorem 10]. In contrast, we have the following result for  $\text{poly-}\mathcal{CF}$  groups.

**Proposition 4.7.** *The restricted standard wreath product  $\mathbb{Z} \wr \mathbb{Z}$  is not  $\text{poly-}\mathcal{CF}$ .*

*Proof.* Since  $\mathbb{Z} \wr \mathbb{Z}$  contains free abelian subgroups of rank  $k$  for all  $k \in \mathbb{N}$ , this follows immediately from Lemma 4.6 and the fact that the  $\text{poly-}\mathcal{CF}$  groups are closed under taking finitely generated subgroups.  $\square$

A further result on wreath products will be useful when we come to consider metabelian groups. It is our first application of Theorem 3.12.

**Proposition 4.8.** *For any  $p \in \mathbb{N} \setminus \{1\}$ , the restricted standard wreath product  $C_p \wr \mathbb{Z}$  is not  $\text{poly-}\mathcal{CF}$ .*

*Proof.* Let  $G = \langle b \rangle \wr \langle a \rangle = C_p \wr \mathbb{Z}$ , with  $p > 1$  and let  $A$  and  $B$  be the inverses of  $a$  and  $b$  respectively. For  $k \in \mathbb{N}$ , let  $W_k = (A^* b a^*)^k (A^* B a^*)^k$  and let  $M_k$  be the sublanguage of  $W_k$  consisting of all those words

$$w = (A^{m_1} b a^{n_1}) \dots (A^{m_k} b a^{n_k}) (A^{m_{k+1}} B a^{n_{k+1}}) \dots (A^{m_{2k}} B a^{n_{2k}})$$

satisfying the following: (i)  $m_i = n_i$  for all  $i$ ; (ii)  $n_i < m_{i+1}$  for  $i \notin \{k, 2k\}$ . Each of (i) and (ii) can be checked by a pushdown automaton, so  $M_k$  is the intersection of two context-free languages and the regular language  $W_k$  and is thus  $2\text{-}\mathcal{CF}$ .

Now let  $L_k = W(G, \{a, b\}) \cap M_k$ . Then  $L_k$  consists of all words of the form

$$b^{a^{m_1}} \dots b^{a^{m_k}} B^{a^{m_{2k+1}}} \dots B^{a^{m_{2k}}} =_G 1,$$

with  $m_i \in \mathbb{N}_0$  for all  $i$ , and  $m_i < m_{i+1}$  for  $i \notin \{k, 2k\}$ . Since the conjugates of  $b$  in such a word are all distinct, for each  $1 \leq i \leq k$  we must have some

$1 \leq j \leq k$  such that  $m_{k+j} = m_i$ . But since  $m_i < m_{i+1}$  and  $m_{k+i} < m_{k+i+1}$  for all  $1 \leq i \leq k-1$ , this means  $m_i = m_{k+i}$  for all  $1 \leq i \leq k-1$ .

When we take  $\Phi(L_k)$ , we can ignore the  $b$ 's and  $B$ 's, since these would contribute nothing to the aspects of the structure of the resulting subset of  $\mathbb{N}_0^{6k}$  that interest us. For our purposes it is equivalent and more straightforward to consider  $\Phi(L_k)$  as a subset of  $\mathbb{N}_0^{4k}$ , thus:

$$\Phi(L_k) = \{(m_1, m_1, \dots, m_k, m_k, m_1, m_1, \dots, m_k, m_k) \mid m_i \in \mathbb{N}_0, m_i < m_{i+1}\}.$$

We see that  $\Phi(L_k)$  is a  $k$ -dimensional subset of the set  $S^{(2,k)}$  studied in Section 3.3. Thus  $\Phi(L_k)$  cannot be expressed as an intersection of  $k-1$  stratified semilinear sets, by Corollary 3.11. Hence  $L_k$  is not  $(k-1)$ - $\mathcal{CF}$ , by Corollary 2.7. Since  $L_k$  is the intersection of  $W(G)$  with a 2- $\mathcal{CF}$  language, this implies that  $W(G)$  is not  $(k-3)$ - $\mathcal{CF}$  for any  $k \in \mathbb{N}$  and so  $G$  is not poly- $\mathcal{CF}$ .  $\square$

### 4.3 The groups $G(\mathbf{c})$

The groups  $G(\mathbf{c})$  were defined in [3] and play an important role in the main results of that paper, which we shall be applying in order to prove certain cases of Conjecture 4.2.

For  $\mathbf{c} = (c_0, \dots, c_s) \in \mathbb{Z}^{s+1}$  with  $s \geq 1$ ,  $c_0, c_s \neq 0$  and  $\gcd(c_0, \dots, c_s) = 1$ , the group  $G(\mathbf{c})$  is defined by the presentation  $\langle a, b \mid \mathcal{R}_{\mathbf{c}} \rangle$ , where

$$\mathcal{R}_{\mathbf{c}} = \left\{ [b, b^{a^i}] \ (i \in \mathbb{Z}), \ b^{c_0} (b^a)^{c_1} \dots (b^{a^s})^{c_s} \right\}.$$

We call such groups *Gc-groups*, and when we refer to the Gc-group  $G(\mathbf{c}) = \langle x, y \rangle$ , we assume that  $\mathbf{c} \in \mathbb{Z}^{s+1}$  satisfies the above conditions, and that  $x$  replaces  $a$  and  $y$  replaces  $b$  in the above definition of  $G(\mathbf{c})$ . Note that here we depart from our usual convention of denoting the  $i$ -th component of  $\mathbf{c}$  by  $\mathbf{c}(i)$ , as it makes the notation more pleasant. A Gc-group is called *proper* if it is not virtually abelian.

As an example, if  $\mathbf{c} = (-m, 1)$  then  $G(\mathbf{c}) = \text{BS}(1, m)$ ; so the soluble Baumslag-Solitar groups are all Gc-groups.

The main result in this section will be that a Gc-group is poly- $\mathcal{CF}$  if and only if it is virtually abelian. We simplify the notation by setting  $b_i = b^{a^i}$  for all  $i \in \mathbb{Z}$ , and  $B = \langle b_i \mid i \in \mathbb{Z} \rangle$ . Since  $B$  is an abelian normal subgroup of

$G(\mathbf{c})$  and  $G(\mathbf{c})/B \cong \langle a \rangle$ , we see that Gc-groups have derived length at most 2.

**Lemma 4.9.** *Let  $G = G(\mathbf{c})$  be a Gc-group with  $|c_0| = |c_s| = 1$ . Then  $G$  is polycyclic.*

*Proof.* The relation  $b_0^{\pm 1} b_1^{c_1} \dots b_{s-1}^{c_{s-1}} b_s^{\pm 1} = 1$  implies that  $b_0, b_{s+1} \in \langle b_1, \dots, b_s \rangle$  and hence  $b_i \in \langle b_1, \dots, b_s \rangle$  for all  $i$ . Hence  $B = \langle b_1, \dots, b_s \rangle$ ; so  $G \triangleright B \triangleright \{1\}$  is a normal series for  $G$  with finitely generated abelian factors and  $G$  is polycyclic.  $\square$

Unsurprisingly, different elements of  $\mathbb{Z}^{s+1}$  can produce isomorphic Gc-groups:

**Lemma 4.10.** *Let  $G = G(\mathbf{c})$ , where  $\mathbf{c} = (c_0, \dots, c_s)$  and let  $\mathbf{c}' = (c_s, c_{s-1}, \dots, c_0)$ . Then  $G(\mathbf{c}) \cong G(\mathbf{c}')$ .*

*Proof.* Let  $G = G(\mathbf{c}) = \langle a, b \rangle$  and let  $x = a^{-1}$  and  $y = b_s$ . Then  $y^{x^i} = b_{s-i}$  for  $i \in \mathbb{Z}$ , so  $b_0^{c_0} b_1^{c_1} \dots b_s^{c_s} = y^{c_s} (y^x)^{c_{s-1}} \dots (y^{x^s})^{c_0}$ . Hence  $G(\mathbf{c}) \cong \langle x, y \rangle = G(\mathbf{c}')$ .  $\square$

The following proposition, proved in [3, Proposition 2.4], gives a useful embedding of a Gc-group in a semidirect product  $\mathbb{Q}^s \rtimes \mathbb{Z}$ .

**Proposition 4.11.** *Let  $G = G(\mathbf{c})$  be a Gc-group. Let  $\{x_1, \dots, x_s\}$  be a basis for  $\mathbb{Q}^s$  over  $\mathbb{Q}$  (the rationals under addition), and let  $\mathbb{Z} = \langle y \rangle$ . Let  $Q = \mathbb{Q}^s \rtimes \mathbb{Z}$ , with the action of  $y$  on  $\mathbb{Q}^s$  being given by the (columns of the) matrix*

$$A(\mathbf{c}) = \begin{pmatrix} 0 & \dots & 0 & -c_0/c_s \\ & & & -c_1/c_s \\ & & & \cdot \\ & I_{s-1} & & \cdot \\ & & & \cdot \\ & & & -c_{s-1}/c_s \end{pmatrix}.$$

*Then  $G$  is isomorphic to the subgroup  $\langle x_1, y \rangle$  of  $Q$ .*

Next, we give a lemma about powers of the matrix  $A(\mathbf{c})$  defined in the previous proposition.

Let  $p$  be a prime. The  $p$ -adic valuation  $v_p : \mathbb{Q} \rightarrow \mathbb{Z} \cup \{\infty\}$  is given by

- $v_p(0) = \infty$ ;
- $v_p(m/n) = d_m - d_n$  for  $m, n \in \mathbb{Z}$ ,  $n \neq 0$ , where  $d_k := \max\{i \in \mathbb{N}_0 \mid p^i \mid k\}$  for all  $k \in \mathbb{Z}$ .

We shall be concerned with powers of a prime occuring in the denominator of various rational numbers. Therefore, rather than  $v_p$ , we shall always be using  $-v_p$ , which, because of the frequency of its occurrence, we shall denote by  $\bar{v}_p$ . Note that if  $\bar{v}_p(a) < \bar{v}_p(b)$ , then  $\bar{v}_p(a+b) = \bar{v}_p(b)$ .

The lemma is stated in slightly more generality than we require, as it is just as easy to prove the more general result.

**Lemma 4.12.** *Let  $M$  be a matrix of the form*

$$\begin{pmatrix} 0 \dots 0 & a_1 \\ & a_2 \\ & \cdot \\ I_{s-1} & \cdot \\ & \cdot \\ & a_s \end{pmatrix},$$

where all  $a_i \in \mathbb{Q}$  and at least one  $a_i \notin \mathbb{Z}$ . Write  $M^k = (m_{ij}^{(k)})$  for  $k \in \mathbb{N}$ . Then there exist  $N \in \{1, \dots, s\}$  and a prime  $p$  such that, for every  $k \in \mathbb{N}$ , there exists some  $i_k \leq ks$  with  $\bar{v}_p(m_{N s}^{(i_k)}) \geq k$ .

*Proof.* Choose some  $a_j \notin \mathbb{Z}$ , and let  $p$  be a prime such that  $\bar{v}_p(a_j) > 0$ . Let  $n = \max\{\bar{v}_p(a_i) \mid 1 \leq i \leq s\}$  and let  $N = \max\{i \mid \bar{v}_p(a_i) = n\}$ . For  $k \in \mathbb{N}$ , denote the entry in the  $N$ -th row and  $s$ -th column of  $M^k$  by  $m_k$ .

Note that for  $k \geq 2$  and  $1 \leq i \leq s-1$ , the  $i$ -th column of  $M^k$  is the same as the  $(i+1)$ -th column of  $M^{k-1}$ . Thus the  $N$ -th row of  $M^k$  is  $(\epsilon_1, \dots, \epsilon_{s-k}, m_1, \dots, m_k)$  if  $k < s$ , with  $\epsilon_i \in \{0, 1\}$ ; and  $(m_{k-s+1}, \dots, m_{k-1}, m_k)$  if  $k \geq s$ . For convenience, rename  $\epsilon_1, \dots, \epsilon_{s-k}$  as  $m_{k-s+1}, \dots, m_0$ , so that we can write the  $N$ -th row of  $M^k$  in the second form in both cases. Notice that  $m_k$  is in the  $N, s-i$  position in  $M^{k+i}$ . In particular, we have  $m_k$  in the  $N, N$  position of  $M^{k+s-N}$  for all  $k \in \mathbb{N}$ .

For  $k \in \mathbb{N}$ , define  $i_k$  to be the minimal natural number such that  $\bar{v}_p(m_{i_k}) \geq k$  if such a number exists, or  $\infty$  otherwise. To begin with, we have  $i_1 = 1$ ,

since  $\bar{v}_p(m_1) = \bar{v}_p(a_N) = n \geq 1$ . We shall show by induction on  $k$  that  $i_k \leq ks$  for all  $k \in \mathbb{N}$ , hence proving the lemma.

Fix  $k \in \mathbb{N}$  and suppose that  $i_k \leq ks$ . Let  $j_k = i_k + s - N$  and consider  $M^{j_k}$ . The  $N$ -th row of this matrix is  $(m_{j_k-s+1}, \dots, m_{i_k}, \dots, m_{j_k-1}, m_{j_k})$ . Note that  $\bar{v}_p(m_{i_k}) \geq k$  and  $\bar{v}_p(m_i) < k$  for  $1 \leq i < i_k$ , by the minimality of  $i_k$ . For  $i \leq 0$ , we have  $m_i \in \{0, 1\}$  and so  $\bar{v}_p(m_i) \in \{0, -\infty\}$ . Thus  $\bar{v}_p(m_i) < k$  for all  $i < i_k$ . Note also that  $j_k + 1 = i_k + s - N + 1 \leq ks + s = (k+1)s$ .

We may assume that  $\bar{v}_p(m_i) \leq k$  for all  $i \leq j_k$ , since otherwise we would have  $i_{k+1} \leq j_k < (k+1)s$  and we would be done. Now

$$\begin{aligned} m_{j_k+1} &= (m_{j_k-s+1}, \dots, m_{i_k}, \dots, m_{j_k}) \cdot (a_1, \dots, a_N, \dots, a_s) \\ &= \sum_{i=1}^s m_{j_k-s+i} a_i = \sum_{i=1}^s m_{i_k-N+i} a_i. \end{aligned} \quad (*)$$

We have  $\bar{v}_p(m_{i_k-N+i} a_i) = \bar{v}_p(m_{i_k-N+i}) + \bar{v}_p(a_i)$  for  $1 \leq i \leq s$ . In particular,  $\bar{v}_p(m_{i_k} a_N) = \bar{v}_p(m_{i_k}) + n \geq k + n$ .

By the maximality of  $N$ , we have  $\bar{v}_p(a_i) < n$  for all  $i > N$ . Since also  $\bar{v}_p(m_{i_k-N+i}) < k$  for  $i < N$ , we thus have  $\bar{v}_p(m_{i_k-N+i} a_i) < k + n$  for  $i \neq N$ . So the  $N$ -th term of  $(*)$  has strictly greater negative  $p$ -adic value than the other terms and hence

$$\bar{v}_p(m_{j_k+1}) = \bar{v}_p(m_{i_k} a_N) \geq k + n \geq k + 1,$$

therefore  $i_{k+1} \leq j_k + 1 \leq (k+1)s$ , as required.  $\square$

We are now ready to prove the main result of this section.

**Proposition 4.13.** *A Gc-group is poly- $\mathcal{CF}$  or co $\mathcal{CF}$  if and only if it is virtually abelian.*

*Proof.* Let  $G = G(\mathbf{c})$  be a proper Gc-group with  $\mathbf{c} \in \mathbb{Z}^{s+1}$ . If  $|c_0| = |c_s| = 1$ , then  $G$  is polycyclic and hence not poly- $\mathcal{CF}$  by Proposition 4.3. Hence if  $|c_s| = 1$ , we may assume  $|c_0| \neq 1$ . By Lemma 4.10,  $G$  is isomorphic to  $G(\mathbf{c}')$ , where  $\mathbf{c}' = (c_s, c_{s-1}, \dots, c_0)$ . Thus we may assume that  $|c_s| \neq 1$ .

By Lemma 4.11, we can identify  $G$  with the subgroup  $\langle x_1, y \rangle$  of  $Q = \mathbb{Q}^s \rtimes \mathbb{Z}$ , where  $\{x_1, \dots, x_s\}$  is a basis for  $\mathbb{Q}^s$  over  $\mathbb{Q}$ ,  $\mathbb{Z} = \langle y \rangle$ , and  $y$  acts on  $\mathbb{Q}^s$  by the matrix  $A(\mathbf{c})$  given in the lemma. Let  $M = A(\mathbf{c})$  and use the notation of

Lemma 4.12 for entries of  $M^k$ . Since  $|c_s| \neq 1$  and  $\gcd(c_0, \dots, c_s) = 1$ , some  $c_i/c_s$  for  $0 \leq i \leq s-1$  is not an integer. Thus  $M$  satisfies the hypothesis of Lemma 4.12. Hence there exist  $I \in \{1, \dots, s\}$  and a prime  $p$  such that, for every  $k \in \mathbb{N}$ , there exists some  $\iota_k \leq ks$  such that  $\bar{v}_p(m_{Is}^{(\iota_k)})$  is at least  $k$ .

For  $k \in \mathbb{N}$ , let

$$\ell_k = \min \left\{ \ell \in \mathbb{N} \mid \ell m_{is}^{(k)} \in \mathbb{Z} \ (1 \leq i \leq s) \right\}.$$

This is the smallest nonnegative integer  $\ell$  such that the final column of  $\ell M^k$  has all integer entries. We are especially interested in the matrices  $M^{\iota_k}$ , and so it will be convenient to set  $\lambda_k = \ell_{\iota_k}$ . Since  $\bar{v}_p(m_{Is}^{(\iota_k)}) \geq k$ , we have  $\lambda_k \geq p^k$  for all  $k \in \mathbb{N}$ .

We can take an increasing sequence of natural numbers  $n_1, n_2, \dots$  such that, for all  $i \in \{1, \dots, s\}$ , the entries  $m_{is}^{(\iota_{n_k})}$  are either nonnegative for all  $k \in \mathbb{N}$ , or negative for all  $k \in \mathbb{N}$ . In the first case we say that  $i$  is of Type 1, while in the second case  $i$  is of Type 2.

We are now ready to define a bounded sublanguage of  $W(G)$  which we can show to be not poly- $\mathcal{CF}$  using Corollary 3.3. Let  $X = \{x_1, \dots, x_s, y\}$  and consider the intersection of  $W(G, X)$  with the bounded context-free language

$$L' = \cup_{k \in \mathbb{N}_0} (y^{-1})^k x_s^* y^k (x_1^{\epsilon_1})^* (x_2^{\epsilon_2})^* \dots (x_s^{\epsilon_s})^*,$$

where  $\epsilon_i = (-1)^j$  if  $i$  is of Type  $j$ . Let  $L = \Phi(W(G, X) \cap L')$ .

The final column of  $M^k$  represents the action of  $y^k$  on  $x_s$ . Specifically,

$$x_s^{y^k} = x_1^{m_{1s}^{(k)}} \dots x_I^{m_{Is}^{(k)}} \dots x_s^{m_{ss}^{(k)}}.$$

For  $\lambda \in \mathbb{Z}$  and  $k \in \mathbb{N}$ , the element  $\left((x_s^\lambda)^{y^k}\right)^{-1}$  of  $G$  can be expressed as a word in  $(x_1^{\epsilon_1})^* (x_2^{\epsilon_2})^* \dots (x_s^{\epsilon_s})^*$  if and only if  $\ell_k \mid \lambda$ . For all  $k \in \mathbb{N}$ , we thus have  $(\iota_k, \lambda, \iota_k; \mathbf{v}) \in L$ , where  $\mathbf{v} \in \mathbb{N}_0^s$ , if and only if  $\ell_{\iota_k} = \lambda_k \mid \lambda$  and  $\mathbf{v}(i) = \lambda |m_{is}^{(\iota_k)}|$  for  $1 \leq i \leq s$ .

Let  $\tau$  be the permutation  $(2, 3)$ . Then for all  $k \in \mathbb{N}$ , we have  $(\iota_k, \iota_k; \mathbf{v}) \in \tau(L)$ , where  $\mathbf{v} \in \mathbb{N}_0^{s+1}$ , if and only if  $\lambda_k \mid \mathbf{v}(1)$  and  $\mathbf{v}(i+1) = \mathbf{v}(1) |m_{is}^{(\iota_k)}|$  for  $1 \leq i \leq s$ .

For  $k \in \mathbb{N}$ , let  $\mathbf{a}_k = ((\iota_{n_k}, \iota_{n_k}))$  and let  $\mathbf{b}_k \in \mathbb{N}_0^{s+1}$  with  $\mathbf{b}_k(1) = \lambda_{n_k}$  and  $\mathbf{b}_k(i+1) = \lambda_{n_k} |m_{is}^{(\iota_{n_k})}|$  for  $1 \leq i \leq s$ . So  $(\mathbf{a}_k; \mathbf{b}) \in \tau(L)$  if and only if  $\mathbf{b}$  is a nonnegative integer multiple of  $\mathbf{b}_k$ .

For any  $t \in \mathbb{N}$ , there exists  $N \in \mathbb{N}$  such that, for all  $k \geq N$ ,

$$t\sigma(\mathbf{a}_k) = 2t\iota_{n_k} \leq 2tsn_k < p^{n_k} \leq \lambda_{n_k} = \mathbf{b}_k(1).$$

Thus, for any  $k \geq N$ ,  $\mathbf{a}_k$  satisfies the first two conditions of Proposition 3.2 with respect to  $t$ . We can take  $k$  such that  $n_k \geq t$ . For any two distinct  $\mathbf{b}$  and  $\mathbf{b}'$  such that  $(\mathbf{a}; \mathbf{b}), (\mathbf{a}; \mathbf{b}') \in \tau(L)$ , there are distinct  $\lambda_1, \lambda_2 \in \mathbb{N}_0$  such that

$$\begin{aligned} |\mathbf{b}(1) - \mathbf{b}'(1)| &= |\lambda_1 \mathbf{b}_k(1) - \lambda_2 \mathbf{b}_k(1)| \\ &= |\lambda_1 - \lambda_2| \lambda_{n_k} \geq p^{n_k} \geq p^t. \end{aligned}$$

Since  $f(t) = p^t$  is an unbounded function, this shows that  $\mathbf{a}_k$  also satisfies the third condition of Proposition 3.2 with respect to  $t$ . Thus  $\tau(L)$  is not a semilinear set and so  $W(G, X) \cap L'$  is neither poly- $\mathcal{CF}$  nor co $\mathcal{CF}$ , by Corollary 3.3. Since  $L'$  is context-free, this implies that  $W(G, X)$  is neither poly- $\mathcal{CF}$  nor co $\mathcal{CF}$ .  $\square$

## 5 Soluble poly- $\mathcal{CF}$ groups

In the case of soluble groups, Conjecture 4.2 simplifies to

**Conjecture 5.1.** *A finitely generated soluble group is poly- $\mathcal{CF}$  if and only if it is virtually abelian.*

Using Theorem 2.12 and the fact that the class of poly- $\mathcal{CF}$  groups is closed under taking finitely generated subgroups (Proposition 2.5), we can make some progress towards resolving Conjecture 5.1.

**Theorem 5.2.** *If  $G$  is a finitely generated poly- $\mathcal{CF}$  soluble group, then one of the following must hold:*

- (i)  $G$  is virtually abelian; or (possibly)
- (ii)  $G$  has a finitely generated subgroup  $H$  with an infinite normal torsion subgroup  $U$  such that  $H/U$  is either free abelian or isomorphic to a proper Gc-group.

*The second case does not occur if  $G$  is metabelian or torsion-free.*



*Proof.* By Theorem 2.12, if  $G$  is a finitely generated soluble group which does not satisfy (i) or (ii), then  $G$  has a subgroup isomorphic to  $\mathbb{Z}^\infty$  or a proper Gc-group.

If  $G$  has a  $\mathbb{Z}^\infty$  subgroup, then  $G$  has free abelian subgroups of rank  $k$  for all  $k \in \mathbb{N}$  and so is not poly- $\mathcal{CF}$  by Lemma 4.6. If  $G$  contains a proper Gc-group, then  $G$  is not poly- $\mathcal{CF}$  by Proposition 4.13.

If  $G$  is torsion-free, then by definition  $G$  has no non-trivial torsion subgroups. If  $G$  is metabelian, then the subgroup  $H$  in the second case can be taken to be  $C_p \wr \mathbb{Z}$  for some prime  $p$ , and hence  $G$  not poly- $\mathcal{CF}$  by Proposition 4.8.  $\square$

We conjecture that the second case does not occur at all, but have been unable to prove this so far.

In order to complete the proof of Conjecture 5.1, we need only show that a finitely generated soluble group  $G$  having an infinite torsion subgroup  $U$  such that  $G/U$  is either free abelian or isomorphic to a proper Gc-group is not poly- $\mathcal{CF}$ .

One way of approaching this which looks promising would be to show that a poly- $\mathcal{CF}$  group cannot have an infinite torsion subgroup. We know that context-free groups cannot have infinite torsion subgroups, because they are virtually free. Actually, we conjecture something stronger, which again is true in the case of context-free groups.

**Conjecture 5.3.** *If a group  $G$  is poly- $\mathcal{CF}$ , then  $G$  does not have arbitrarily large finite subgroups.*

So far, the author's approaches towards this conjecture, from the perspective of automata theory, have not succeeded. It may be that an approach using grammars would be more fruitful.

## 5.1 An example of the undetermined case

We give a proof of non-poly-context-freeness in a specific example of the second case of Theorem 5.2.

If  $\langle X \mid R \rangle$  is a group presentation, we denote the abelianisation of the group with this presentation by  $\text{Ab} \langle X \mid R \rangle$ . This enables us to write shorter pre-

sentations for abelian groups, by omitting the commutators of generators from the relator set. We call such a presentation an *abelian presentation*.

**Proposition 5.4.** *Let  $p$  be a prime and let  $G$  be the group given by the following presentation.*

$$\langle a, b_i \ (i \in \mathbb{Z}), c_j \ (j > 0) \mid \begin{array}{l} b_i^a = b_{i+1} \ (i \in \mathbb{Z}), [b_i, b_{i+j}] = c_j \ (i \in \mathbb{Z}, j > 0), \\ b_i^p = c_j^p = 1 \ (i \in \mathbb{Z}, j > 0), c_j \text{ central } (j > 0) \end{array} \rangle.$$

*Then  $G$  has derived length 3 and satisfies (ii) of Theorem 5.2, and is not poly-CF.*

*Proof.* In this proof, we shall always assume that the indices on the right hand side of a presentation run over all available values (specified on the left hand side). This prevents the presentations from becoming too cluttered. With this convention, the presentation for  $G$  is simplified to

$$\langle b_i \ (i \in \mathbb{Z}), c_j \ (j > 0) \mid b_i^a = b_{i+1}, [b_i, b_{i+j}] = c_j, b_i^p = c_j^p = 1, c_j \text{ central} \rangle.$$

Let  $H$  be the group defined by the subpresentation

$$\langle b_i \ (i \in \mathbb{Z}), c_j \ (j > 0) \mid [b_i, b_{i+j}] = c_j, b_i^p = c_j^p = 1, c_j \text{ central} \rangle.$$

Then  $a$  acts on  $H$  by conjugation as an automorphism of infinite order, so  $G \cong H \rtimes \langle a \rangle$  and  $G/H \cong \mathbb{Z}$ . Thus  $G$  satisfies the second case of Theorem 5.2, with  $U = H$ . Since  $G \triangleright H \triangleright \langle c_j \ (j > 0) \rangle \triangleright \{1\}$  is a normal series for  $G$  with abelian factors,  $G$  has derived length at most 3.

By standard results on ‘Darstellungsgruppen’ (covering groups) in [18, Chapter V.23], in the group  $E_n$  given by the presentation

$$\langle b_i \ (-n \leq i \leq n), c_{ij} \ (-n \leq i < j \leq n) \mid [b_i, b_j] = c_{ij}, b_i^p = c_{ij}^p = 1, c_{ij} \text{ central} \rangle,$$

the subgroup generated by all the  $c_{ij}$  (which is  $E'_n$ ) has the abelian presentation  $\text{Ab} \langle c_{ij} \ (-n \leq i < j \leq n) \mid c_{ij}^p \rangle$ .

Let  $E$  be the union of the ascending sequence of groups  $E_1, E_2, \dots$ . Then  $E' = \cup_{i \in \mathbb{N}} E'_i$ , with presentation  $\text{Ab} \langle c_{ij} \ (i, j \in \mathbb{Z}, i < j) \mid c_{ij}^p \rangle$ . Our subgroup  $H$  of  $G$  is obtained from  $E$  by quotienting out the subgroup  $N := \langle c_{0,j-i} c_{ij}^{-1} \mid i < j \rangle$  and setting  $c_j = c_{0j}$  for all  $j > 0$ . The subgroup of  $H$

generated by all the  $c_j$  is isomorphic to  $E'/N$ , and thus has abelian presentation

$$\text{Ab} \langle c_j \ (j > 0) \mid c_j^p \rangle.$$

In particular, all  $c_j$  are non-trivial and so  $H$  is not abelian, and therefore  $G$  has derived length 3.

Let  $b = b_0$ ,  $B = B_0$  and let  $M_k$  be the sublanguage of

$$W_k = (BA^*Ba^*bA^*ba^*)^k (BA^*ba^*bA^*Ba^*)^k$$

consisting of all those words

$$(BA^{m_1}Ba^{n_1}bA^{\mu_1}ba^{\nu_1}) \dots (BA^{m_k}Ba^{n_k}bA^{\mu_k}ba^{\nu_k})(BA^{m_{k+1}}ba^{n_{k+1}}bA^{\mu_{k+1}}Ba^{\nu_{k+1}}) \dots (BA^{m_{2k}}ba^{n_{2k}}bA^{\mu_{2k}}Ba^{\nu_{2k}})$$

such that: (i)  $m_i = n_i = \mu_i = \nu_i$  for all  $i$ ; (ii)  $m_i < m_{i+1}$  for  $i \notin \{k, 2k\}$ . The first condition can be checked by two pushdown automata, one checking that  $m_i = n_i$  and  $\mu_i = \nu_i$  for all  $i$ , and the other checking that  $m_i = \mu_i$  for all  $i$ . The second condition can be checked by a single pushdown automaton. Thus  $M_k$  is  $3\text{-}\mathcal{CF}$ .

A word in  $M_k$  is equal in  $G$  to

$$[b, b_{m_1}] \dots [b, b_{m_k}][b, B_{m_{k+1}}] \dots [b, B_{m_{2k}}] = c_{m_1} \dots c_{m_k} (c_{m_{k+1}})^{-1} \dots (c_{m_{2k}})^{-1},$$

with  $m_i < m_{i+1}$  and  $m_{k+i} < m_{k+i+1}$  for  $1 \leq i \leq k-1$ .

Let  $L_k = \Phi(W(G) \cap M_k)$ . As in the proof of Proposition 4.8, we can ignore the  $b$ 's and  $B$ 's and take  $L_k$  to be a subset of  $\mathbb{N}_0^{8k}$ . Since the  $c_{m_i}$  are distinct for  $1 \leq i \leq k$  and

$$\langle c_j \mid j > 0 \rangle = \text{Ab} \langle c_j \ (j > 0) \mid c_j^p \ (j > 0) \rangle,$$

the only way that a word in  $M_k$  can be in  $W(G)$  is if some  $m_{k+j} = m_i$  for each  $1 \leq i \leq k$ . But since  $m_i < m_{i+1}$  and  $m_{k+i} < m_{k+i+1}$  for  $1 \leq i \leq k-1$ , this implies that  $m_i = m_{k+i}$  for  $1 \leq i \leq k$  and so  $L_k$  is the set of all  $8k$ -tuples of the form

$$(m_1, m_1, m_1, m_1, \dots, m_k, m_k, m_k, m_k, m_1, m_1, m_1, m_1, \dots, m_k, m_k, m_k, m_k),$$

with  $m_i \in \mathbb{N}_0$ , and  $m_i < m_{i+1}$  for  $1 \leq i \leq k-1$ . Thus  $L_k$  is a  $k$ -dimensional linear subset of the set  $S^{(4,k)}$  introduced in Section 3.3, and is therefore not

an intersection of  $k - 1$  stratified semilinear sets, by Corollary 3.11. By Corollary 2.7, this means that  $W(G) \cap M_k$  is not  $(k - 1)$ - $\mathcal{CF}$ . Since  $M_k$  is  $3$ - $\mathcal{CF}$ , this implies that  $W(G)$  is not  $(k - 4)$ - $\mathcal{CF}$  for any  $k \in \mathbb{N}$ . Hence  $G$  is not poly- $\mathcal{CF}$ .  $\square$

Quotienting out a proper subgroup of  $\langle c_j \ (j > 0) \rangle$  in the group  $G$  in Proposition 5.4 results in another group of derived length 3 satisfying (ii) of Theorem 5.2. We do not know how to show that such quotients are not poly- $\mathcal{CF}$  except in some very specific cases.

**Acknowledgements** I am immensely grateful to my Ph.D. supervisor, Derek Holt, for many helpful and inspiring discussions and suggestions. This research was supported by a Vice Chancellor’s Scholarship from the University of Warwick.

## References

- [1] V. A. Anisimov, The group languages, *Kibernetika* **4** (1971), 18–24.
- [2] T. Brough, *Groups with poly-context-free word problem*, Ph.D. thesis, University of Warwick, 2010.
- [3] T. Brough and D. F. Holt, Finitely generated soluble groups and their subgroups, preprint, 2011. <http://arxiv.org/abs/1009.4149>.
- [4] M. Dunwoody, The accessibility of finitely presented groups, *Invent. Math.* **81** (1985), 449–457.
- [5] S. Ginsburg, S. A. Greibach and M. A. Harrison, One-way stack automata, *J. ACM* **14** (1967), 389–418.
- [6] S. Ginsburg, *The mathematical theory of context-free languages*, McGraw-Hill, 1966.
- [7] S. Ginsburg and E. H. Spanier, Bounded ALGOL-like languages, *Trans. Amer. Math. Soc.* **113** (1964), 333–368.
- [8] S. Ginsburg and E. H. Spanier, Semigroups, Presburger formulas, and languages, *Pacif. J. Math.* **16** (1966), 285–296.
- [9] I. Gorun, A heirarchy of context-sensitive languages, *Lect. Notes Comput. Sc.* **45** (1976), 299–303.

- [10] T. Herbst, On a subclass of context-free groups, *Theoret. Inform. Appl.* **25** (1991), 255–272.
- [11] D.F. Holt, S. Rees, C.E. Röver and R.M. Thomas, Groups with context-free co-word problem, *J. London Math. Soc.* (2) **71** (2005), 643–657.
- [12] D. F. Holt and S. Rees, Solving the word problem in real time, *J. London Math. Soc.* (2) **63** (2001), no.3, 623–639.
- [13] D. F. Holt and C. E. Röver, On real-time word problems, *J. London Math. Soc.* (2) **67** (2003), no.2, 289–301.
- [14] D. F. Holt and C. E. Röver, Groups with indexed co-word problem, *Internat. J. Algebra Comput.* **16** (2006), no.5, 985–1014.
- [15] D. F. Holt, M. D. Owens and R. M. Thomas, Groups and semigroups with a one-counter word problem, *J. Aust. Math. Soc.* **85** (2008) 197–209.
- [16] D. F. Holt, S. Rees and M. Shapiro, Groups that do and do not have growing context-sensitive word problem, *Internat. J. Algebra Comput.* (2008), no.7, 1179–1191.
- [17] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, 1979.
- [18] B. Huppert, *Endliche Gruppen I*, Springer-Verlag, 1967.
- [19] M. Kambites and F. Otto, Church-Rosser groups and growing context-sensitive groups, *J. Autom. Lang. Comb.* **13** (2008), no.3-4, 249–267.
- [20] J. Lehnert and P. Schweitzer, The co-word problem for the Higman-Thompson group is context-free, *Bull. Lond. Math. Soc.* **39** (2007), no.2, 235–241.
- [21] L. Y. Liu and P. Weiner, An infinite heirarchy of intersections of context-free languages, *Math. Systems Theory* **7** (1973), 185–192.
- [22] R. C. Lyndon and P. E. Schupp, *Combinatorial Group Theory*, Springer-Verlag, 1977.
- [23] D. E. Muller and P. E. Schupp, Groups, the theory of ends, and context-free languages, *J. Comput. System Sci.* **26** (1983), 295–310.

- [24] D. E. Muller and P. E. Schupp, The theory of ends, pushdown automata, and second-order logic, *Theoret. Comput. Sci.* **37** (1985), 51–75.
- [25] R. J. Parikh, Language-generating devices, *MIT Res. Lab. Electron. Quart. Prog. Rep.* **60** (1961), 199–212.
- [26] M. Shapiro, A note on context-sensitive languages and word problems, *Internat. J. Algebra Comput.* **4** (1994), no.4, 493–497.
- [27] D. Wotschke, The Boolean closures of deterministic and nondeterministic context-free languages, *Lect. Notes Comput. Sc.* **1** (1973), 113–121.